PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [6] : C07K 16/46, 16/00, A61K 39/395, G01N 33/577

A1

(11) International Publication Number: WO 97/14719

(43) International Publication Date: 24 April 1997 (24.04.97)

(21) International Application Number: PCT/EP96/03605

(22) International Filing Date: 14 August 1996 (14.08.96)

(30) Priority Data:
95307332.7 16 October 1995 (16.10.95) GB

(71) Applicant (*for all designated States except AU BB CA GB IE LK MN MW NZ SD US*): UNILEVER N.V. [NL/NL]; Weena 455, NL-3013 AL Rotterdam (NL).

(71) Applicant (*for AU BB CA GB IE LK MN MW NZ SD only*): UNILEVER PLC [GB/GB]; Unilever House, Blackfriars, London EC4 4BQ (GB).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DAVIS, Paul, James [GB/GB]; The Hawthorns, Pavenham Road, Felmersham, Bedfordshire MK43 7EX (GB). VAN DER LOGT, Cornelis, Paul, Erik [NL/GB]; 1 Bluebell Rise (Peverel Manor Estate), Rushden, Northamptonshire MK43 7EX (GB). VERHOEIJEN, Martine, Elisa [BE/GB]; 1 Tintagel Close (Manor Farm Estate), Rushden, Northamptonshire NN10 0TU (GB). WILSON, Steve [GB/GB]; 3 Aldenham Close (Goldington), Bedford MK41 0FQ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: A BIFUNCTIONAL OR BIVALENT ANTIBODY FRAGMENT ANALOGUE

(57) Abstract

The invention relates to a bispecific or bivalent antibody fragment analogue comprising a binding complex containing two polypeptide chains, whereby one polypeptide chain comprises two heavy chain variable domains (V_H) in series and the other polypeptide chain comprises two light chain variable domains (V_L) in series, the binding complex further containing two pairs of variable domains (V_H-A//V_L-A and V_H-B//V_L-B). The two V_H's and/or the two V_L's are connected directly or via an intermediate peptide linker. Also a production method for such antibody fragment analogues is disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

Title: A **bifunctional** or **bivalent** antibody fragment analogue

The invention relates to new **bispecific** or **bivalent** antibody fragment analogues, a process for preparing such antibody fragment analogues and various uses of such antibody fragment analogues.

Background of the invention and prior art

1. Antibody structure

Antibody molecules typically are Y-shaped molecules whose basic unit consist of four polypeptides, two identical heavy chains and two identical light chains, which are covalently linked together by disulfide bonds. Each of these chains is folded in discrete domains. The C-terminal regions of both heavy and light chains are conserved in sequence and are called the constant regions, also known as C-domains. The N-terminal regions, also known as V-domains, are variable in sequence and are responsible for the antibody specificity. The antibody specifically recognizes and binds to an antigen mainly through six short complementarity-determining regions located in their V-domains (see Figure 1).

In this specification abbreviations are used having the following meaning.

C-domain:		Constant domain
V-domain:		Variable domain
V_L	:	Variable domain of the light chain
V_H	:	Variable domain of the heavy chain
Fv	:	dual chain antibody fragment containing both a V_H and a V_L
scFv	:	single-chain Fv (V_H and V_L genetically linked either directly or via a peptide linker)
CDR	:	Complementarity-determining region
ELISA	:	Enzyme Linked Immuno Sorbent Assay
PCR	:	Polymerase Chain Reaction

IPTG	:	IsoPropyl-ß-ThioGalactopyranoside
PBS	:	Phosphate Buffered Saline
PBST	:	Phosphate Buffered Saline with 0.15% Tween
TMB	:	3,3',5,5'-TetraMethylBenzidine

It is generally known that proteolytic digestion of an antibody with papain yields three fragments. The fragment containing the CH_2 and CH_3 domains of the two heavy chains connected by the complete hinge (see Figure 1) crystallises very easily and was therefore called Fc fragment. The two other fragments are identical and were called Fab fragments, as they contained the antigen-binding site. Digestion with pepsin is such that the two Fab's remain connected via the hinge, forming only two fragments: Fc' and Fab_2.
The Fv is the smallest unit of an antibody which still contains the complete binding site (see Figure 1) and full antigen binding activity. It consists of only the V-domains of the heavy and light chains thus forming a small, heterodimeric variable fragment or Fv. Fv's have a molecular weight of about 25 kD, which is only one sixth of the parent whole antibody (in the case of an IgG). Previously Fv's were only available by proteolysis in a select number of cases (Givol, 1991). The production of Fv's can now be achieved more routinely using genetic engineering methods through cloning and expressing DNA encoding only the V-domains of the antibody of interest. Smaller fragments, such as individual V-domains (Domain Antibodies or dABs, Ward *et al.*, 1989), and even individual CDR's (Williams *et al.*, 1989; Taub *et al.*, 1989) were shown to retain the binding characteristics of the parent antibody. However, this is not achievable on a routine basis: most naturally occurring antibodies need both a V_H and a V_L to retain full immunoreactivity. For example, in the case of V_H D1.3 (Ward *et al.*, 1989), although it still binds hen egg lysozyme (HEL) with an affinity close to that of the parent antibody, it was shown that loss of

specificity was observed in that it can no longer distinguish turkey lysozyme from HEL, whereas the Fv can (Berry and Davies, 1992). Although murine dABs can be obtained more routinely from spleen libraries (Ward *et al.*, 1989), the approach is unsustainable because of the many problems associated with their production and physical behaviour: expression is extremely poor, affinity tends to be low, stability and solubility in water is low, and non-specific binding is usually very high. According to the literature a possible explanation of these undesirable characteristics is the exposure of the hydrophobic residues which are normally buried in the V_H-V_L interface. The exposed hydrophobic patches are thought to contribute to aggregation of the protein inside the cells and/or in the culture medium, leading to poor expression and/or poor solubility (Anthony *et al.*, 1992; Ward *et al.*, 1989). The hydrophobic patches can also explain the high non-specific binding described by Berry and Davies, 1992. These problems clearly limit the usefulness of these molecules.

Most of the Camelid antibodies appear to be an exception to this rule in that they only need one V-domain, namely V_H, to specifically and effectively bind an antigen (Hamers-Castermans *et al.*, 1993). In addition, preliminary data indicate that they seem not to suffer from the disadvantages of mouse dABs, as these camelid antibodies or fragments thereof are soluble and have been shown to express well in yeast and *Aspergillus* moulds. These observations can have important consequences for the production and exploitation of antibody-based products, see patent application WO 94/25591 (UNILEVER *et al.*, first priority date 29.04.93).

2. Production of antibody fragments

Several microbial expression systems have already been developed for producing active antibody fragments, e.g. the production of Fab in various hosts, such as *E. coli* (Better *et al.*, 1988, Skerra and Plückthun, 1988, Carter *et al.*,

1992), yeast (Horwitz *et al.*, 1988), and the filamentous fungus *Trichoderma reesei* (Nyyssönen *et al.*, 1993) has been described. The recombinant protein yields in these alternative systems can be relatively high (1-2 g/l for Fab secreted to the periplasmic space of *E. coli* in high cell density fermentation, see Carter *et al.*, 1992), or at a lower level, e.g. about 0.1 mg/l for Fab in yeast in fermenters (Horwitz *et al.*, 1988), and 150 mg/l for a fusion protein CBHI-Fab and 1 mg/l for Fab in *Trichoderma* in fermenters (Nyyssönen *et al.*, 1993) and such production is very cheap compared to whole antibody production in mammalian cells (hybridoma, myeloma, CHO). Although the latter can give yields of the order of 1 g/l in high cell density fermentation, it is a time-consuming and very expensive manufacturing method resulting in a cost price of about 1000 £/gram of antibody. It was further demonstrated that plants can be used as hosts for the production of both whole antibodies (Hiatt *et al.*, 1989) and scFv's (Owen *et al.*, 1992, Firek *et al.*, 1993), whereby yields of upto 0.5% of the total soluble protein content in tobacco leaves were mentioned.

The fragments can be produced as Fab's or as Fv's, but additionally it has been shown that a V_H and a V_L can be genetically linked in either order by a flexible polypeptide linker, which combination is known as an scFv (Bird *et al.* (1988), Huston *et al.* (1988), and granted patent EP-B-0281604 (GENEX/ENZON LABS INC.; first priority date 02-09-1986).

3. Bivalent and bispecific antibodies and antibody fragments

The antibody fragments Fab, Fv and scFv differ from whole antibodies in that the antibody fragments carry only a single antigen-binding site. Recombinant fragments with two binding sites have been made in several ways, for example, by chemical cross-linking of cysteine residues introduced at the C-terminus of the V_H of an Fv (Cumber *et al.*, 1992),

or at the C-terminus of the V_L of an scFv (Pack and Plückthun, 1992), or through the hinge cysteine residues of Fab's (Carter *et al.*, 1992). Another approach to produce bivalent antibody fragments is described by Kostelny *et al.* (1992) and Pack and Plückthun (1992) and is based on the inclusion of a C-terminal peptide that promotes dimerization.

When two different specificities are desired, one can generate bispecific antibody fragments. The traditional approach to generate bispecific whole antibodies was to fuse two hybridoma cell lines each producing an antibody having the desired specificity. Because of the random association of immunoglobulin heavy and light chains, these hybrid hybridomas produce a mixture of up to 10 different heavy and light chain combinations, only one of which is the bispecific antibody (Milstein and Cuello, 1983). Therefore, these bispecific antibodies have to be purified with cumbersome procedures, which considerably decrease the yield of the desired product.

Alternative approaches include in-vitro linking of two antigen specificities by chemical cross-linking of cysteine residues either in the hinge or via a genetically introduced C-terminal Cys as described above. An improvement of such in vitro assembly was achieved by using recombinant fusions of Fab's with peptides that promote formation of heterodimers (Kostelny *et al.*, 1992). However, the yield of bispecific product in these methods is far less than 100%.

A more efficient approach to produce bivalent or bispecific antibody fragments, not involving in vitro chemical assembly steps, was described by Holliger *et al.* (1993). This approach takes advantage of the observation that scFv's secreted from bacteria are often present as both monomers and dimers. This observation suggested that the V_H and V_L of different chains can pair, thus forming dimers and larger complexes. The dimeric antibody fragments, also named "diabodies" by Hollinger *et al.*, in fact are small

bivalent antibody fragments that assembled in vivo. By linking the V_H and V_L of two different antibodies 1 and 2, to form "cross-over" chains V_H1V_L2 and V_H2-V_L1 (see Figure 2B), the dimerisation process was shown to reassemble both antigen-binding sites. The affinity of the two binding sites was shown to be equal to the starting scFv's, or even to be 10-fold increased when the polypeptide linker covalently linking V_H and V_L was removed, thus generating two proteins each consisting of a V_H directly and covalently linked to a V_L not pairing with the V_H (see Figure 2C). This strategy of producing bispecific antibody fragments was also described in several patent applications. Patent application WO 94/09131 (SCOTGEN LTD; priority date 15.10.92) relates to a bispecific binding protein in which the binding domains are derived from both a V_H and a V_L region either present at two chains or linked in an scFv, whereas other fused antibody domains, e.g. C-terminal constant domains, are used to stabilise the dimeric constructs. Patent application WO 94/13804 (CAMBRIDGE ANTIBODY TECHNOLOGY / MEDICAL RESEARCH COUNCIL; first priority date 04.12.92) relates to a polypeptide containing a V_H and a V_L which are incapable of associating with each other, whereby the V-domains can be connected with or without a linker.

Mallender and Voss, 1994 (also described in patent application WO 94/13806; DOW CHEMICAL CO; priority date 11.12.92) reported the in vivo production of a single-chain bispecific antibody fragment in *E. coli*. The bispecificity of the bivalent protein was based on two previously produced monovalent scFv molecules possessing distinct specificities, being linked together at the genetic level by a flexible polypeptide linker. The thus formed V_H1-linker-V_L1-linker-V_H2-linker-V_L2 fragment (see Figure 2A) was shown to contain both antigen binding specificities 1 and 2. (1= anti-fluorescein, 2= anti-single-stranded DNA).

Traditionally, whenever single-chain antibody fragments are referred to, a single molecule consisting of one heavy chain linked to one (corresponding) light chain in the presence or absence of a polypeptide linker is implicated. When making bivalent or bispecific antibody fragments through the 'diabody' approach (Holliger *et al.*, (1993) and patent application WO 94/09131) or by the 'double scFv' approach (Mallender and Voss, 1994 and patent application WO 94/13806), again the V_H is linked to a (the corresponding) V_L.

It is realised that claims 32 and 33 of patent application WO 93/11161 (ENZON INC.; priority date 25.11.91) and the corresponding passages in that specification on page 22, lines 1-10 may read on a polypeptide comprising two V_L's fused together via a flexible polypeptide linker, and on a polypeptide comprising two V_H's fused together via a flexible polypeptide linker, respectively. However, no examples were given to substantiate this approach, thus it was in fact a hypothetical possibility instead of an actually produced compound.

A skilled person would not have expected that such approach would be viable for at least three reasons. Firstly, it is widely recognised that immunoglobulin heavy chains (excluding the above described camel immunoglobulins) have very limited solubility and spontaneously precipitate out of aqueous solution when isolated from their light chain partners. Secondly, several groups have shown (Ward *et al.*, 1989, Berry and Davies, 1992, and Anthony *et al.*, 1992) that expression of V_H's in the absence of V_L's is hampered by extremely poor yields of unstable product with many undesirable properties, e.g. non-specific binding. Thirdly in patent application WO 94/13804 it was described on page 31 lines 10-12, that in computer modelling experiments they could not model as heterodimers V_H-V_H and V_L-V_L given the constraints of short linkers.

Thus the simple suggestion given in patent application WO 93/11161 is not an enabling disclosure leading a skilled

person to try with a reasonable expectation of success whether such suggestion would work; therefore, that patent application should not be considered as relevant prior art for the present invention.

Summary of the invention

The present invention provides a **bispecific** or **bivalent** antibody fragment analogue, which comprises a binding complex containing two polypeptide chains, one of which comprises two times a variable domain of a heavy chain (V_H) **in series** and the other comprises two times a variable domain of a light chain (V_L) **in series**.

In one aspect of the invention one chain of the antibody fragment analogue comprises a first V_H (V_H-A) <u>connected to</u> a second V_H (V_H-B) and the other chain comprises a first V_L (V_L-A) <u>connected to</u> a second V_L (V_L-B). In a preferred embodiment of this aspect one chain comprises a first V_H (V_H-A) <u>followed by</u> a second V_H (V_H-B), thus [V_H-A * V_H-B], and the other chain comprises a first V_L (V_L-A) <u>preceded by</u> a second V_L (V_L-B), thus [V_L-B * V_L-A]. For some embodiments of this aspect the two V_H's are directly connected to each other, but for other embodiments of this aspect of the invention the two V_L's are directly connected to each other.

According to another embodiment of this aspect of the invention the two V_H's are connected to each other by a linker and also the two V_L's are connected to each other by a linker. Such a linker usually comprises at least one amino acid residue.

According to a special embodiment of this aspect of the invention one chain comprises a first V_H (V_H-A) <u>followed by</u> a second V_H (V_H-B), thus [V_H-A * V_H-B], and the other chain comprises a first V_L (V_L-A) <u>followed by</u> a second V_L (V_L-B), thus [V_L-A * V_L-B], and in which the two V_H's are connected to each other by a linker and also the two V_L's are connected to each other by a linker, whereas each linker comprises at least 10 amino acid residues.

According to the above aspect of the invention with A being different from B there are provided **bispecific** antibody fragment analogues.

According to another aspect of the invention the specificities A and B are the same resulting in a **bivalent** antibody fragment.

According to a further aspect of the invention the bispecific or bivalent antibody fragment analogues can be used in a diagnostic technique or for immunoassays, in a purification method, for therapy, or in other methods in which immunoglobulins or fragments thereof are used. Such uses are well-known in the art.

The invention also provides a process for producing the antibody fragments of the invention in that a host is transformed by incorporating into that host a DNA encoding the two V_H's with or without a connecting linker and a DNA encoding the two V_L's with or without a connecting linker. Preferably the two DNA's are placed in a dicistronic arrangement.

It is also possible that the two linked V_H's and the two linked V_L's are produced separately by different hosts, after which the linked V_H's produced by one host can be combined with the linked V_L's produced by the other host. The hosts can be selected from the group consisting of prokaryotic bacteria of which examples are Gram-negative bacteria, e.g. *E. coli*, and Gram-positive bacteria, e.g. *B. subtilis* or lactic acid bacteria, lower eukaryotes examples of which are yeasts, e.g. belonging to the genera *Saccharomyces*, *Kluyveromyces*, or *Trichoderma*, moulds, e.g. belonging to the genera *Aspergillus* and *Neurospora*, and higher eukaryotes, examples of which are plants, e.g. tobacco, and animal cells, examples of which are myeloma cells and CHO cells. The techniques to transform a host by genetic engineering methods in order to have a desirable polypeptide produced by such host are well-known to persons skilled in the art as is evident from the literature

mentioned above under the heading "Background of the invention and prior art".

Brief description of the drawings

Figure 1 depicts in schematic form the structure of a typical antibody (immunoglobulin) molecule.

Figure 2 shows a schematic representation of published arrangements of heavy and light chain V-domain gene fragments that have been proven to produce bispecific antibody fragments.

Figure 3 shows in diagrammatic form the suggested arrangement of the V-domains of a double head antibody fragment according to the invention with the V-domains in the following order: V_HA-V_HB + V_LB-V_LA.

Figure 4 shows the nucleotide sequence of the *Eco*RI-*Hin*dIII insert of pUR.4124 containing DNA (see SEQ ID NO: 1) encoding V_LLys-Linker-V_HLys (see SEQ ID NO: 2).

Figure 5 shows the nucleotide sequence of the *Hin*dIII-*Eco*RI insert of plasmid Fv.3418 (see SEQ ID NO: 3) containing DNA encoding pelB leader-V_H3418 (see SEQ ID NO: 4) and DNA encoding pelB leader-V_L3418 (see SEQ ID NO: 5).

Figure 6 shows the nucleotide sequence of the *Hin*dIII-*Eco*RI insert of plasmid Fv.4715-myc (see SEQ ID NO: 6) containing DNA encoding pelB leader-V_H4715 (see SEQ ID NO: 7) and DNA encoding pelB leader-V_L4715-Myc tag (see SEQ ID NO: 8).

Figure 7 shows the nucleotide sequence of the *Hin*dIII-*Eco*RI insert of scFv.4715-myc containing DNA (see SEQ ID NO: 9) encoding pelB leader-V_H4715-Linker-V_L4715-Myc tag (see SEQ ID NO: 10).

Figure 8 a/b shows the nucleotide sequence of the *Hin*dIII-*Eco*RI insert of pGOSA.E (see SEQ ID NO: 11) containing DNA encoding pelB leader-V_H4715-

Linker-V_L3418 (see SEQ ID NO: 12) and DNA encoding pelB leader-V_L3418-Linker-V_H4715 (see SEQ ID NO: 13).

Figure 9 gives an overview of the oligonucleotides and their positions in pGOSA.E that can be used to replace V-domain gene fragments.

Figure 10 illustrates the amino acid sequence of the V_H-V_H and V_L-V_L domain junctions in fusion polypeptides GOSA.E (see amino acids 114-145 in SEQ ID NO: 12 and amino acids 102-128 in SEQ ID NO: 13), GOSA.V (see SEQ ID NO: 30 and amino acids 102-128 in SEQ ID NO: 13), GOSA.S (see amino acids 114-145 in SEQ ID NO: 12 and SEQ ID NO: 31) and GOSA.T (see SEQ ID NO: 30 and SEQ ID NO: 31).

Figure 11 shows the specificity of *Streptococcus* binding of scFv.4715-myc.

Figure 12 shows the specificity of glucose oxidase targeting onto the surface of various Streptococcus strains by GOSA.E.

Figure 13 shows the specificity of glucose oxidase targeting onto the surface of various Streptococcus strains by GOSA.V.

Figure 14 shows the specificity of glucose oxidase targeting onto the surface of various *Streptococcus* strains by GOSA.S.

Figure 15 shows the specificity of glucose oxidase targeting onto the surface of various *Streptococcus* strains by GOSA.T.

Figure 16 shows the results of an ELISA. Individual fractions of a gelfiltration experiment using partially purified GOSA.E as feedstock were tested for glucose oxidase and *Streptococcus sanguis* bispecific binding activity.

Figure 17 shows the results of an ELISA. Individual fractions of a gelfiltration experiment using partially purified GOSA.V as feedstock were

tested for glucose oxidase and *Streptococcus sanguis* bispecific binding activity.

Figure 18 shows the results of an ELISA. Individual fractions of a gelfiltration experiment using partially purified GOSA.S as feedstock were tested for glucose oxidase and *Streptococcus sanguis* bispecific binding activity.

Figure 19 shows the results of an ELISA. Individual fractions of a gelfiltration experiment using partially purified GOSA.T as feedstock were tested for glucose oxidase and *Streptococcus sanguis* bispecific binding activity.

Figure 20 shows the source of fragment PCR.I *Bst*EII/*Sac*I

Figure 21 shows the source of fragment PCR.II *Sfi*I/*Eco*RI

Figure 22 shows the source of fragment PCR.III *Nhe*I/*Sac*I

Figure 23 shows the source of fragment PCR.IV *Xho*I/*Eco*RI

Figure 24 shows the source of fragment PCR.V *Sal*I/*Eco*RI

Figure 25 shows the source of fragment PCR.VI *Sfi*I/*Nhe*I

Figure 26 shows the source of fragment PCR.VII *Bst*EII/*Nhe*I

Figure 27 shows the source of fragment PCR.VIII *Xho*I/*Eco*RI

Figure 28 shows the source of fragment PCR.IX *Bst*EII/*Nhe*I

Figure 29 shows the source of fragment PCR.X *Pst*I/*Eco*RI

Figure 30 shows the construction of plasmid pGOSA.A

Figure 31 shows the construction of plasmid pGOSA.B

Figure 32 shows the construction of plasmid pGOSA.C

Figure 33 shows the construction of plasmid pGOSA.D

Figure 34 shows the construction of plasmid pGOSA.E

Figure 35 shows the construction of plasmid pGOSA.V

Figure 36 shows the construction of plasmid pGOSA.S

Figure 37 shows the construction of plasmid pGOSA.T

Figure 38 a/b shows the construction of plasmid pGOSA.G

Figure 39 shows the construction of plasmid pGOSA.J

Figure 40 shows the construction of plasmid pGOSA.Z

Figure 41 shows the construction of plasmid pGOSA.AA

Figure 42	shows the construction of plasmid pGOSA.AB
Figure 43	shows the construction of plasmid pGOSA.L
Figure 44	shows the construction of plasmid pGOSA.Y
Figure 45	shows the construction of plasmid pGOSA.X
Figure 46	shows the construction of plasmid pGOSA.AC
Figure 47	shows the construction of plasmid pGOSA.AD.

Table 1 shows the nucleotide sequence of the oligonucleotides used to produce the constructs described in this specification. Restriction sites encoded by these primers are underlined.

Table 2 gives an overview of all GOSA constructs described in this specification.
Table 2A describes intermediate constructs that were not further tested.
Table 2B describes the dicistronic constructs.
Table 2C describes the monocistronic constructs.

Detailed description of the invention

In this specification the construction of an antibody fragment analogue consisting of a two chain protein complex is described, in which one of the chains consists of two heavy chain V-domains and the other chain consists of the two corresponding light chain V-domains in either order. The variable domains are linked either directly or through a polypeptide linker. Subsequent molecular modelling of this combination suggested that the protein chains could fold such that both binding sites are fully accessible, provided that the connecting linkers are kept long enough to span 30 to 35 Å.
Whereas in patent application WO 93/11161 it is explicitly described that for the above described bispecific complexes two flexible polypeptide linkers in the self assembling complex are required, the present invention illustrated here describes in particular the construction of a two chain protein complex containing only one linker or no

linkers at all. The latter antibody fragment analogue thus consists of a two chain protein complex containing one polypeptide chain comprising heavy chain V-domains fused directly together and another polypeptide chain comprising the corresponding light chain V-domains fused together, both fusions in the absence of linkers. But also two chain protein complexes in which each chain comprises a linker between the two variable domains can be used as antibody fragment analogues according to the invention as described below with construct pGOSA.E. However, the two chain complexes containing only one linker or no linker at all are preferred. The abbreviation GOSA used in this specification relates to a combination of glucose oxidase and *Streptococcus sanguis*.

In this specification evidence is provided that these antibody fragment analogues ("double heads") contain both antigen binding specificities of the Fv's used to generate these bispecific antibody fragments. It is exemplified that these type of constructs according to the invention can be used to target the enzyme glucose oxidase to whole bacteria, using antibody fragments derived from hybridomas expressing antibodies directed against these antigens.

The present invention is now described by reference to some specific examples, which are included for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLES

General experimental

Strains, Plasmids and Media

All cloning steps were performed in *E. coli* JM109 (*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17(r_K, m_K^+), *rel*A1, *sup*E44, Δ(*lac*-proAB), [F', *tra*D36, *pro*AB, *lac*I^{q}ZΔM15]. *E. coli* cultures were grown in 2xTY medium (16 g tryptone, 10 g yeast extract, 5 g NaCl per litre H_2O), where indicated

supplemented with 2% glucose and/or 100 μg/ml ampicillin. Transformations were plated out on SOBAG plates (20 g tryptone, 5 g yeast extract, 15 g agar, 0.5 g NaCl per litre H_2O plus 10 mM $MgCl_2$, 2% glucose, 100 μg/ml ampicillin) The expression vectors used are derivatives of pUC19. The oligonucleotide primers used in the PCR reactions were synthesized on an Applied Biosystems 381A DNA Synthesiser by the phosphoramidite method.

Expression of GOSA constructs

Colonies from freshly transformed JM109 plated onto SOBAG plates were used to inoculate 2xTY medium supplemented with 100 μg/ml ampicillin, 2% glucose. Cultures were shaken at 37°C to an OD_{600} in the range of 0.5 to 1.0. Cells were pelleted by centrifugation and the supernatant was removed. The pelleted cells were resuspended in 2xTY medium with 100 μg/ml ampicillin, 1 mM IPTG, and grown for a further 18 hours at 25°C. Cells were pelleted by centrifugation and the supernatant, containing the secreted chains, used directly in an ELISA. The proteins in the periplasm of the pelleted cells were extracted by resuspending the cell pellet in 1/20 of the original culture volume of lysis buffer (20% sucrose, 200 mM Tris-HCl pH 7.5, 1 mM EDTA, 500 μg/ml lysozyme). After incubation at 25°C for 20 minutes an equal volume of H_2O was added and the incubation was continued for another 20 minutes. The suspension was spun at 10.000 g for 15 minutes and the supernatant containing the periplasmic proteins was used directly in an ELISA.

ELISA

96 well ELISA plates (Greiner HC plates) were activated overnight at 37°C with 200 μl/well of an 1/10 dilution of an over night culture of *Streptococcus* cells in 0.05 M sodium carbonate buffer at pH=9.5. Following one wash with PBST, the antigen sensitised plates were pre-blocked for 1 hour at 37°C with 200 μl/well blocking buffer (2% BSA, 0.15% Tween in PBS). Samples containing 50 μl blocking

buffer plus 50 μl culture supernatants or periplasmic cell extracts (neat or diluted with PBS) were added to the *Streptococcus* sensitised plate and incubated for 2 hours at 37°C. Following 4 washes with PBS-T, 100 μl of blocking buffer containing glucose oxidase (50 μg/ml) was added to every well. After incubation at 37°C for 1 hour unbound glucose oxidase was removed by 4 washes with PBS-T. Bound glucose oxidase was detected by adding 100 μl substrate to each well (70 mM Na-citrate, 320 mM Na-phosphate, 27 mg/ml glucose, 0.5 μg/ml HRP, 100 μg/ml TMB). The colour reaction was stopped after 1 hour by the addition of 35 μl 2 M HCl and the A450 was measured (compare Figures 11/15).

Affinity purification of GOSA antibody fragments

GOSA.E, GOSA.V, GOSA.S and GOSA.T were partially purified by affinity chromatography. 100 ml periplasmic extract of each of these constructs was loaded onto a Glucose-oxidase-Sepharose column (CNBr-Sepharose, Pharmacia) prepared according to the manufacturer's instructions. After extensive washes with PBS the bound GOSA antibody fragments were eluted in 0.1M glycine buffer at pH=2.8. The fractions were neutralised with Tris and analysed by polyacrylamide gel electrophoresis followed by silver staining and tested for the presence of double head activity.

EXAMPLE 1. Construction of the pGOSA double head expression vectors

In this Example the construction of a two chain protein complex is described, in which one of the chains consists of two heavy chain V-domains and the other chain consists of the two corresponding light chain V-domains. The variable domains are linked either directly or through a polypeptide linker. The expression vectors used are derivatives of a pUC19 derived plasmid containing a *Hin*dIII-*Eco*RI fragment that in the case of plasmid

scFv.4715-myc contains a DNA fragment encoding one pelB signal sequence fused to the N-terminus of the V_H that is directly linked to the corresponding V_L of the antibody through a connecting flexible peptide linker, $(Gly_4Ser)_3$ (present in SEQ ID NO: 2 as amino acids 109-123 and in SEQ ID NO: 10 as amino acids 121-135), thus generating a single-chain molecule (see Figure 7).
In the dual-chain Fv and the pGOSA expression vectors, the DNA fragments encoding both the V_H and V_L of the antibody are preceded by a ribosome binding site and a DNA sequence encoding the pelB signal sequence in an artificial dicistronic operon under the control of a single inducible promoter (see Figures 5, 6 and 8). Expression of these constructs is driven by the inducible lacZ promoter. The nucleotide sequence of the *Hin*dIII-*Eco*RI inserts of the plasmids pUR.4124, Fv.3418, Fv.4715-myc and scFv.4715-myc constructs used for the generation of the bispecific antibody fragments are given in Figures 4-7, respectively.
Moreover, a culture of *E. coli* cells harbouring plasmid scFv.4715-myc and a culture of *E. coli* cells harbouring plasmid Fv.3418 were deposited under the Budapest Treaty at the National Collection of Type Cultures (Central Public Health Laboratory) in London (United Kingdom) with deposition numbers NCTC 12916 and NCTC 12915, respectively.
In agreement with Rule 28 (4) EPC, or a similar arrangement for a State not being a Contracting State of the EPC, it is hereby requested that a sample of such deposit, when requested, will be submitted to an expert only.
The construction of pGOSA.E (see Figure 8 for the *Hin*dIII-*Eco*RI insert of pUC19) involved several cloning steps. The appropriate restriction sites in the various domains were introduced by PCR directed mutagenesis using the oligonucleotides listed in Table 1 below. The pGOSA.E derivatives pGOSA.V, pGOSA.S and pGOSA.T with only one or no linker sequence are derived from the pGOSA.E construct by removing the linker sequences by means of PCR directed mutagenesis with oligonucleotides listed in Table 1 below.

Table 1.

DBL.1	5'-CAC CAT CTC CAG AGA CAA TGG CAA G-3'	(=SEQ ID NO: 14)
DBL.2	5'-GAG CGC GAG CTC GGC CGA ACC GGC C[1]GA TCC GCC ACC GCC AGA GCC-3'	(=SEQ ID NO: 15)
DBL.3	5'-CAG GAT CCG GCC GGT TCG GCC[1] CAG GTC CAG CTG CAA CAG TCA GGA-3'	(=SEQ ID NO: 16)
DBL.4	5'-CTA CAT GAA TTC[2] GCT AGC[3] TTA TTA TGA GGA GAC GGT GAC GGT GGT CCC TTG GC-3'	(=SEQ ID NO: 17)
DBL.5	5'-TAA TAA GCT AGC[3] GGA GCT GCA TGC AAA TTC TAT TTC-3'	(= SEQ ID NO: 18)
DBL.6	5'-ACC AAG CTC GAG[4] ATC AAA CGG GG-3'	(= SEQ ID NO: 19)
DBL.7	5'-AAT GTC GAA TTC[2] GTC GAC[5] TCC GCC ACC GCC AGA GCC-3'	(= SEQ ID NO: 20)
DBL.8	5'-ATT GGA GTC GAC[5] ATC GAA CTC ACT CAG TCT CCA TTC TCC-3'	(= SEQ ID NO: 21)
DBL.9	5'-TGA AGT GAA TTC[2] GCG GCC GC[6]T TAT TAC CGT TTG ATT TCG AGC TTG GTC CC-3'	(= SEQ ID NO: 22)
DBL.10	5'-CGA ATT CGG TCA CC[8]G TCT CCT CAC AGG TCC AGT TGC AAC AG-3'	(= SEQ ID NO: 23)
DBL.11	5'-CGA ATT CTC GAG[4] ATC AAA CGG GAC ATC GAA CTC ACT CAG TCT CC-3'	(= SEQ ID NO: 24)
DBL.12	5'-CGA ATT CGG TCA CC[8]G TCT CCT CAC AGG TGC AGT TGC AGG AG-3'	(= SEQ ID NO: 25)
PCR.51	5'-AGG T(C/G)(A/C) A(C/A)C TGC AG[7](C/G) AGT C(A/T)G G-3'	(= SEQ ID NO: 26)
PCR.89	5'-TGA GGA GAC GGT GAC C[8]GT GGT CCC TTG GCC CC-3'	(= SEQ ID NO: 27)
PCR.90	5'-GAC ATT GAG CTC[9] ACC CAG TCT CCA-3'	(= SEQ ID NO: 28)
PCR.116	5'-GTT AGA TCT CGA G[4]CT TGG TCC C-3'	(= SEQ ID NO: 29)

1=*Sfi*I, 2=*Eco*RI, 3=*Nhe*I, 4=*Xho*I, 5=*Sal*I, 6=*Not*I, 7=*Pst*I, 8=*Bst*EII, 9=*Sac*I

These three constructs lack some of the restriction sites at the new joining points. The V_HA-V_HB gene fragment without a linker lacks the 5' V_HB *Sfi*I site. The V_LB-V_LA gene fragment without a linker lacks the 5' V_LA *Sal*I site. The position of the oligonucleotides in the pGOSA constructs given in Table 1 are shown in Figure 9. The pGOSA expression vectors and the oligonucleotides in Table 1 have been designed to enable most specificities to be cloned into the pGOSA constructs. Figure 10 shows the amino acid sequence of the junctions between the V_HA-V_HB and V_LB-V_LA fragments encoded by DNA present in pGOSA.E, pGOSA.V, pGOSA.S and pGOSA.T. A more detailed description of the preparation of pGOSA.E, pGOSA.V, pGOSA.S and pGOSA.T is given in Example 5.

EXAMPLE 2. Bifunctional binding activity of GOSA double heads

In this Example we provide evidence that the above described molecules ("double heads"), i.e. the two chain protein complexes, contain both antigen binding specificities of the Fv's used to generate these multi-functional antibody fragment analogues. Figure 12-15 show that GOSA.E, GOSA.V, GOSA.S and GOSA.T can be used to specifically target the enzyme glucose oxidase to several *Streptococcus sanguis* strains using antibody fragments derived from hybridoma's expressing antibodies directed against these antigens.
Comparison of the binding specificity of the GOSA constructs (see Figures 12-15) and the binding specificity of the scFv.4715-myc (see Figure 11) shows that the fine specificity of the anti-*Streptococcus sanguis* scFv.4715 is preserved in the GOSA "double heads".

EXAMPLE 3. FPLC analysis of GOSA double heads

Partially purified GOSA.E, GOSA.V, GOSA.S and GOSA.T samples (estimated to be 50-80% pure by polyacrylamide gel electrophoresis) were analysed on a Pharmacia FPLC Superose 12 column. The analysis was performed using PBS at a flow rate of 0.3 ml/minute. Eluate was monitored at 280 nm and 0.3 ml fractions were collected and analysed by ELISA. Usually GOSA.E, GOSA.V, GOSA.S and GOSA.T samples only gave one GOSA double head activity peak as determined by ELISA (see Figures 16-19). The position of this peak in the elution pattern indicated that the molecular weight of the GOSA double head is 40-50 kD. Since this molecular weight corresponds to the expected molecular weight of the V_H2 + V_L2 double head dimer, the conclusion is justified that GOSA.E, GOSA.V, GOSA.S and GOSA.T are primarily produced as dimeric molecules. Occasionally an activity peak with an apparent molecular weight of ~200 kD was observed (see Figure 16). The presence of Glucose Oxidase activity in these fractions (data not shown) indicate that these fractions contain GOSA double head complexed with glucose oxidase that was eluted with the GOSA sample from the glucose oxidase-sepharose affinity matrix.

EXAMPLE 4. Production of other double heads

The methods described in the previous Examples were used to produce other double heads, which also appeared to be active against the antigens for which they were developed. These other double heads had the following specificities:
anti-*S. sanguis* / anti-beta-HCG,
anti-*S. sanguis* / anti-urease,
anti-*S. sanguis* / anti-hen-egg-lysozyme,
anti-beta-HCG / anti-hen-egg-lysozyme,
anti-hen-egg-lysozyme /anti-glucose oxidase,
anti-huIgG / anti-glucose oxidase,

anti-urease / anti-glucose oxidase,
anti-lacto-peroxidase / anti-glucose oxidase,
anti-alpha-HCG / anti-glucose oxidase, and
anti-reactive-Red-6 / anti-glucose oxidase.

EXAMPLE 5. Detailed description of the preparation of intermediate constructs pGOSA.A, pGOSA.B pGOSA.C and pGOSA.D and their use for the preparation of plasmid pGOSA.E and its derivatives pGOSA.V, pGOSA.S and pGOSA.T

Oligonucleotides and PCR

The primary structures of the oligonucleotide primers used in the construction of the bispecific 'pGOSA' constructs are shown in Table 1 above. Reaction mixture used for amplification of DNA fragments were 10 mM Tris-HCl, pH 8.3, 2.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin (w/v), 0.1% Triton X-100, 400 mM of each dNTP, 5.0 units of Vent DNA polymerase (New England Biolabs), 100 ng of template DNA, and 500 ng of each primer (for 100 μl reactions). Reaction conditions were: 94°C for 4 minutes, followed by 33 cycles of each 1 minute at 94°C, 1 minute at 55°C, and 1 minute 72°C.

Plasmid DNA\Vector\Insert preparation and ligation\transformation.

Plasmid DNA was prepared using the 'Qiagen P-100 Midi-DNA Preparation' system. Vectors and inserts were prepared by digestion of 10 μg (for vector preparation) or 20 μg (for insert preparation) with the specified restriction endonucleases under appropriate conditions (buffers and temperatures as specified by suppliers). Modification of the DNA ends with Klenow DNA polymerase and dephosphorylation with Calf Intestine Phosphorylase were performed according to the manufacturers instructions. Vector DNAs and inserts were separated through agarose gel electrophoresis and purified with DEAE-membranes NA45

(Schleicher & Schuell) as described by Maniatis *et al.* Ligations were performed in 20 μl volumes containing 30 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 300-400 ng vector DNA, 100-200 ng insert DNA and 1 Weiss unit T_4 DNA ligase. After ligation for 2-4 h at room temperature, $CaCl_2$ competent *E. coli* JM109 (Maniatis) were transformed using 7.5 μl ligation reaction. The transformation mixtures were plated onto SOBAG plates and grown overnight at 37°C. Correct clones were identified by restriction analysis and verified by automated dideoxy sequencing (Applied Biosystems).

Restriction digestion of PCR products

Following amplification each reaction was checked for the presence of a band of the appropriate size by agarose gel electrophoresis. One or two 100 μl PCR reaction mixtures of each of the PCR reactions PCR.I - PCR.X (Figure 20-29), together containing approximately 2-4 μg DNA product were subjected to phenol-chloroform extraction, chloroform extraction and ethanol precipitation. The DNA pellets were washed twice with 70% ethanol and allowed to dry. Next, the PCR products were digested overnight (18 h) in the presence of excess restriction enzyme in the following mixes at the specified temperatures and volumes.

PCR.I:

50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 4 mM spermidine, 0.4 μg/ml BSA, 4 μl (= 40 U) *Sac*I, 4 μl (= 40 U) *Bst*EII, in 100 μl total volume at 37°C.

PCR.II:

10 mM Tris-Acetate pH 7.5, 10 mM $MgAc_2$, 50 mM KAc (1x "One-Phor-All" buffer ex Pharmacia), 4 μl (= 48 U) *Sfi*I, in 50 μl total volume at 50°C under mineral oil. After overnight digestion, PCR.II-*Sfi*I was digested with *Eco*RI (overnight at 37°C) by the addition of 16 μl H_2O, 30 μl 10x "One-Phor-All" buffer (Pharmacia)(100 mM Tris-Acetate pH 7.5, 100 mM $MgAc_2$, 500 mM KAc) and 4 μl (= 40 U) *Eco*RI.

PCR.III:

10 mM Tris-Acetate pH 7.5, 10 mM $MgAc_2$, 50 mM KAc (1x "One-Phor-All" buffer {Pharmacia}), 4 µl (= 40 U) *Nhe*I, 4 µl (= 40 U) *Sac*I, in 100 µl total volume at 37°C.

PCR.IV:

20 mM Tris-Acetate pH 7.5, 20 mM $MgAc_2$, 100 mM KAc (2x "One-Phor-All" buffer {Pharmacia}), 4µl (= 40 U) *Xho*I, 4 µl (= 40 U) *Eco*RI, in 100 µl total volume at 37°C.

PCR.V:

20 mM Tris-Acetate pH 7.5, 20 mM $MgAc_2$, 100 mM KAc (2x "One-Phor-All" buffer {Pharmacia}), 4 µl (= 40 U) *Sal*I, 4 µl (= 40 U) *Eco*RI, in 100 µl total volume at 37°C.

PCR.VI:

10 mM Tris-Acetate pH 7.5, 10 mM $MgAc_2$, 50 mM KAc (1x "One-Phor-All" buffer {Pharmacia}), 4 µl (= 48 U) *Sfi*I, in 50 µl total volume at 50°C under mineral oil. After overnight digestion, PCR.VI-*Sfi*I was digested with *Nhe*I (overnight at 37°C) by the addition of 41 µl H_2O, 5 µl 10x "One-Phor-All" buffer (Pharmacia) (100 mM Tris-Acetate pH 7.5, 100 mM $MgAc_2$, 500 mM KAc) and 4 µl (= 40 U) *Nhe*I.

PCR.VII:

50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 4 mM spermidine, 0.4 µg/ml BSA, 4 µl (= 40 U) *Nhe*I, 4 µl (= 40 U) *Bst*EII, in 100 µl total volume at 37°C.

PCR.VIII:

20 mM Tris-Acetate pH 7.5, 20 mM $MgAc_2$, 100 mM KAc (2x "One-Phor-All" buffer {Pharmacia}), 4 µl (= 40 U) *Eco*RI, in 50 µl total volume at 37°C. After overnight digestion, PCR.VIII-*Eco*RI was digested with *Xho*I (overnight at 37°) by the addition of 46 µl H_2O and 4 µl (= 40 U) *Xho*I.

PCR.IX:

25 mM Tris-Acetate, pH 7.8, 100 mM KAc, 10 mM MgAc, 1mM DTT (1x "Multi-Core" buffer {Promega}), 4 mM spermidine, 0.4 µg/ml BSA, 4 µl (= 40 U) *Nhe*I, 4 µl (= 40 U) *Bst*EII, in 100 µl total volume at 37°C.

PCR.X:
50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 4 mM spermidine, 0.4 μg/ml BSA, 4 μl (= 40 U) *Pst*I, 4 μl (= 40 U) *Eco*RI, in 100 μl total volume at 37°C.

The digested PCR fragments

PCR.I-*Sac*I/*Bst*EII,	PCR.II-*Sfi*I/*Eco*RI,
PCR.III-*Nhe*I/*Sac*I,	PCR.IV-*Xho*I/*Eco*RI,
PCR.V-*Sal*I/*Eco*RI,	PCR.VI-*Sfi*I/*Nhe*I,
PCR.VII-*Bst*EII/*Nhe*I,	PCR.VIII-*Xho*I/*Eco*RI,
PCR.IX-*Bst*EII/*Nhe*I, and	PCR.X-*Pst*I/*Eco*RI

were purified on an 1.2% agarose gel using DEAE-membranes NA45 (Schleicher & Schuell) as described by Maniatis *et al*. The purified fragments were dissolved in H_2O at a concentration of 100-150 ng/μl.

Construction of the pGOSA double head expression vectors. The construction of pGOSA.E (see Figure 8) involved several cloning steps that produced 4 intermediate constructs pGOSA.A to pGOSA.D (see Figure 30-34). The final expression vector pGOSA.E and the oligonucleotides in Table 1 above have been designed to enable most specificities to be cloned into the final pGOSA.E construct (Figure 9). The upstream V_H domain can be replaced by any *Pst*I-*Bst*EII V_H gene fragment obtained with oligonucleotides PCR.51 and PCR.89 (see Table 1 above). The oligonucleotides DBL.3 and DBL.4 (see Table 1 above) were designed to introduce *Sfi*I and *Nhe*I restriction sites in the V_H gene fragments thus allowing cloning of those V_H gene fragments into the *Sfi*I-*Nhe*I sites as the downstream V_H domain. All V_L gene fragments obtained with oligonucleotides PCR.116 and PCR.90 (see Table 1 above) can be cloned into the position of the V_L.3418 gene fragment as a *Sac*I-*Xho*I fragment. A complication here however is the presence of an internal *Sac*I site in the V_H.3418 gene fragment. Oligonucleotides DBL.8 and DBL.9 (see Table 1 above) are designed to allow cloning of V_L gene fragments into the position of the

V_L.4715 gene fragment as a *Sal*I-*Not*I fragment. The pGOSA.E derivatives pGOSA.V, pGOSA.S and pGOSA.T with only one or no linker sequences contain some aberrant restriction sites at the new joining points. The V_HA-V_HB construct without a linker lacks the 5' V_HB *Sfi*I site. The V_HB fragment is cloned into these constructs as a *Bst*EII/*Nhe*I fragment using oligonucleotides DBL.10 or DBL.11 and DBL.4 (see Table 1 above). The V_LB-V_LA construct without a linker lacks the 5' V_LA *Sal*I site. The V_LA fragment is cloned into these constructs as a *Xho*I/*Eco*RI fragment using oligonucleotides DBL.11 and DBL.9 (see Table 1 above).

In the following part of the description the following linkers are mentioned which are also present in the sequence listing:
the $(Gly_4Ser)_3$ linker, present in SEQ ID NO: 2 as amino acids 109-123 and SEQ ID NO: 10 as amino acids 121-135,
the $(Gly_4Ser)_3$AlaGlySerAla linker (= linkerA), present in SEQ ID NO: 12 as amino acids 121-139, and
the $(Gly_4Ser)_2Gly_4$Val linker (= linkerV), present in SEQ ID NO: 13 as amino acids 108-122.

pGOSA.A

This plasmid is derived from both the Fv.4715-myc construct and the scFv.4715-myc construct.
An *Sfi*I restriction site was introduced between the DNA sequence encoding the $(Gly_4Ser)_3$ linker and the gene fragment encoding the V_L of the scFv.4715-myc construct (see Figure 30). This was achieved by replacing the *Bst*EII-*Sac*I fragment of the latter construct by the fragment PCR-I *Bst*EII/*Sac*I (Figure 20) that contains an *Sfi*I site between the DNA encoding the $(Gly_4Ser)_3$ linker and the V_L.4715 gene fragment. The introduction of the *Sfi*I site also introduced 4 additional amino acids (AlaGlySerAla) between the $(Gly_4Ser)_3$ linker and V_L.4715 resulting in a $(Gly_4Ser)_3$AlaGlySerAla linker (linkerA). The oligonucleotides used to produce PCR-I (DBL.1 and DBL.2,

see Table 1 above) were designed to match the sequence of the framework-3 region of V_H.4715 and to prime at the junction of the DNA encoding the $(Gly_4Ser)_3$ linker and the V_L.4715 gene fragment, respectively. Thus pGOSA.A can be indicated as: ***pel*B-V_H4715-linkerA-(*Sfi*I)-V_L4715-myc.**

pGOSA.B

This plasmid is derived from plasmid Fv.3418 (see Figure 31). The *Xho*I-*Eco*RI fragment of plasmid Fv.3418 comprising the 3' end of DNA encoding framework-4 of the V_L including the stop codon was removed and replaced by the fragment PCR-IV *Xho*I/*Eco*RI (Figure 23). The oligonucleotides used to produce PCR-IV (DBL.6 and DBL.7, see Table 1 above) were designed to match the sequence at the junction of the V_L and the $(Gly_4Ser)_3$ linker perfectly (DBL.6), and to be able to prime at the junction of the $(Gly_4Ser)_3$ linker and the V_H in pUR.4124 (DBL.7). DBL.7 removed the *Pst*I site in the V_H (silent mutation) and introduced a *Sal*I restriction site at the junction of the $(Gly_4Ser)_3$ linker and the V_H, thereby replacing the last Ser of the linker by a Val residue resulting in a $(Gly_4Ser)_2Gly_4Val$ linker (linkerV). Thus pGOSA.B can be indicated as:

***pel*B-V_H3418 + *pel*B-V_L3418-linkerV-(*Sal*I-*Eco*RI)**

pGOSA.C

This plasmid contains DNA encoding V_H.4715 linked by the $(Gly_4Ser)_3$AlaGlySerAla linker to V_H.3418 (see Figure 32), thus: ***pel*B-V_H4715-linkerA-V_H3418.**

This construct was obtained by replacing the *Sfi*I-*Eco*RI fragment from pGOSA.A encoding V_L.4715 by the fragment PCR-II *Sfi*I/*Eco*RI containing the V_H.3418 gene (see Figure 21). The oligonucleotides used to produce PCR-II (DBL.3 and DBL.4, see Table 1 above) hybridize in the framework-1 and framework-4 region of the gene encoding V_H.3418, respectively. DBL.3 was designed to remove the *Pst*I restriction site (silent mutation) and to introduce an *Sfi*I restriction site upstream of the V_H gene. DBL.4 destroys

the *Bst*EII restriction site in the framework-4 region and introduces an *Nhe*I restriction site downstream of the stopcodon.

pGOSA.D

This plasmid contains a dicistronic operon comprising the V_H.3418 gene and DNA encoding V_L.3418 linked by the $(Gly_4Ser)_2Gly_4Val$ linker to V_L.4715 (see Figure 33), thus:

***pelB*-V_H3418 + *pelB*-V_L3418-linkerV-V_L4715.**

This construct was obtained by digesting plasmid pGOSA.B with *Sal*I-*Eco*RI and inserting the fragment PCR-V *Sal*I/*Eco*RI (Figure 24) containing the V_L.4715 gene. The oligonucleotides used to obtain PCR-V (DBL.8 and DBL.9, see Table 1 above) were designed to match the nucleotide sequence of the framework-1 and framework-4 regions of the V_L.4715 gene, respectively. DBL.8 removed the *Sac*I site from the framework-1 region (silent mutation) and introduced a *Sal*I restriction site upstream of the V_L.4715 gene. DBL.9 destroyed the *Xho*I restriction site in the framework-4 region of the V_L.4715 gene (silent mutation) and introduced a *Not*I and an *Eco*RI restriction site downstream of the stop codon.

pGOSA.E

This plasmid contains a dicistronic operon comprising DNA encoding V_H.4715 linked by the $(Gly_4Ser)_3$AlaGlySerAla linker to V_H.3418 plus DNA encoding V_L.3418 linked by the $(Gly_4Ser)_2Gly_4Val$ linker to V_L.4715 (see Figure 34), thus:

***pelB*-V_H4715-linkerA-V_H3418 + *pelB*-V_L3418-linkerV-V_L4715.**

Both translational units are preceded by a ribosome binding site and DNA encoding a pelB leader sequence. This plasmid was obtained by a three-point ligation by mixing the vector resulting from pGOSA.D after removal of the V_H3418-encoding *Pst*I-*Sac*I insert with the *Pst*I-*Nhe*I pGOSA.C insert containing V_H.4715 linked to V_H.3418 and the PCR-III *Nhe*I/*Sac*I fragment (see Figure 22). The remaining *Pst*I-*Sac*I pGOSA.D vector contains the 5′ end of the framework-1

region of V_H.3418 upto the *Pst*I restriction site and V_L.3418 linked by the $(Gly_4Ser)_2Gly_4Val$ linker to V_L.4715 starting from the *Sac*I restriction site in V_L.3418. The *Pst*I-*Nhe*I pGOSA.C insert contains V_H.4715 linked by the $(Gly_4Ser)_3$-AlaGlySerAla linker to V_H.3418, starting from the *Pst*I restriction site in the framework-1 region in V_H.4715. The *Nhe*I-*Sac*I PCR-III fragment provides the ribosome binding site and DNA encoding the *pel*B leader sequence for the V_L.3418-$(Gly_4Ser)_2Gly_4Val$-V_L.4715 construct. The oligonucleotides DBL.5 and PCR.116 (see Table 1 above) used to generate PCR-III were designed to match the sequence upstream of the ribosome binding site of V_L.4715 in Fv.4715 and to introduce an *Nhe*I restriction site (DBL.5), and to match the framework-4 region of V_L.3418 (PCR.116).

pGOSA.V

This plasmid is derived from pGOSA.E (see Figure 35) from which the *Bst*EII/*Nhe*I fragment containing DNA encoding linkerA-V_H.3418 was excised and replaced by the fragment PCR-VII *Bst*EII/*Nhe*I containing the V_H.3418 gene (see Figure 26). The resulting plasmid pGOSA.V contains V_H.3418 linked directly to the framework-4 region of V_H.4715, plus V_L.4715 linked by the $(Gly_4Ser)_2Gly_4Val$ linker to the framework-4 region of V_L.3418, thus:

pelB-V_H4715*V_H3418 + pelB-V_L3418-linkerV-V_L4715.

pGOSA.S

This plasmid is derived from pGOSA.E (see Figure 36) from which the $(Gly_4Ser)_2Gly_4Val$-V_L4715 *Xho*I/*Eco*RI fragment was excised and replaced by the fragment PCR-VIII *Xho*I/*Eco*RI which contains V_L.4715 (see Figure 27). The resulting plasmid pGOSA.S contains V_H.4715 linked by the $(Gly_4Ser)_3$-AlaGlySerAla linker to V_H.3418 plus V_L.3418 linked directly to the 5′ end of the framework-1 region of V_L.4715, thus:

pelB-V_H.4715-linkerA-V_H.3418 + pelB-V_L.3418*V_L.4715.

pGOSA.T

This plasmid contains a dicistronic operon consisting of V_H.3418 directly to the framework-4 region of V_H.4715 plus V_L.3418 linked directly to the 5′ end of the framework-1 region of V_L.4715 (see Figure 37). Both transcriptional units are preceded by a ribosome binding site and a *pel*B leader sequence, thus:

_pel_B-V_H.4715*V_H.3418 + _pel_B-V_L.3418*V_L.4715.

This construct was obtained by inserting the *Nhe*I/*Eco*RI fragment of pGOSA.S which contains V_L.3418 linked directly to the 5′ end of the framework-1 region of V_L.4715, into the vector pGOSA.V from which the *Nhe*I/*Eco*RI fragment containing V_L.3418 linked by the (Gly$_4$Ser)$_2$Gly$_4$Val linker to V_L.4715 was removed.

EXAMPLE 6. Detailed description of the preparation of other dicistronic constructs pGOSA.G, and pGOSA.J, pGOSA.Z, pGOSA.AA and pGOSA.AB

pGOSA.G

This plasmid is an intermediate for the synthesis of pGOSA.J. It is derived from pGOSA.E from which the V_H4715 *Pst*I/*Bst*EII fragment has been excised and replaced by the V_H3418 *Pst*I/*Bst*EII fragment (excised from Fv.3418). The resulting plasmid pGOSA.G (see Figure 38) contains two copies of V_H.3418 linked by the (Gly$_4$Ser)$_3$AlaGlySerAla linker, plus V_L.4715 linked by the (Gly$_4$Ser)$_2$Gly$_4$Val linker to the framework-4 region of V_L.3418, thus:

_pel_B-V_H.3418-linkerA-V_H.3418 + _pel_B-V_L.3418-linkerV-V_L.4715.

pGOSA.J

This plasmid contains a dicistronic operon consisting of V_H.3418 linked by the (Gly$_4$Ser)$_3$AlaGlySerAla linker to V_H.4715 plus V_L.3418 linked by the (Gly$_4$Ser)$_2$Gly$_4$Val linker to V_L.4715. Both transcriptional units are preceded by a

ribosome binding site and a *pelB* leader sequence (see Figure 39), thus:

***pelB*-V_H.3418-linkerA-V_H.4715 + *pelB*-V_L.3418-linkerV-V_L.4715.**

This construct was obtained by inserting the fragment PCR-VI *Sfi*I/*Nhe*I which contains V_H4715 (Figure 25), into the vector pGOSA.G from which the *Sfi*I/*Nhe*I V_H3418 fragment was removed.

pGOSA.Z

This plasmid is derived from pGOSA.G from which the $(Gly_4Ser)_3$AlaGlySerAla linker-V_H3418 *Bst*EII/*Nhe*I fragment was excised and replaced by the fragment PCR-IX *Bst*EII/*Nhe*I which contains V_H.4715 (Figure 28). The resulting plasmid pGOSA.Z (see Figure 40) contains V_H.3418 linked directly to the framework-1 region of V_H.4715, plus V_L.4715 linked by the $(Gly_4Ser)_2Gly_4Val$ linker to the framework-4 region of V_L.3418, thus:

***pelB*-V_H.3418*V_H.4715 + *pelB*-V_L.3418-linkerV-V_L.4715.**

pGOSA.AA

This plasmid contains a dicistronic operon consisting of the V_H.3418 linked directly to the 5' end of the framework-1 region of V_H.4715 plus V_L.3418 linked directly to the 5' end of the framework-1 region of V_L.4715. Both transcriptional units are preceded by a ribosome binding site and a *pelB* leader sequence (see Figure 41). This construct was obtained by inserting the *Nhe*I/*Eco*RI fragment of pGOSA.T which contains V_L.3418 linked directly to the 5' end of the framework-1 region of V_L.4715, into the vector pGOSA.Z from which the *Nhe*I/*Eco*RI fragment containing V_L.3418 linked by the $(Gly_4Ser)_2Gly_4Val$ linker to V_L.4715 was removed, thus:

***pelB*-V_H.3418*V_H.4715 + *pelB*-V_L.3418*V_L.4715.**

pGOSA.AB

This plasmid is derived from pGOSA.J by a three point ligation reaction (see Figure 42). The *Sac*I/*Eco*RI insert, containing part of V_H.3418 and the full $(Gly_4Ser)_3$AlaGlySerAla linker-V_H.4715 and the V_L.3418-$(Gly_4Ser)_2Gly_4$Val-V_L.4715 encoding sequences, was removed and replaced by the *Sac*I/*Sac*I pGOSA.J fragment containing the same part of V_H.3418 and the full $(Gly_4Ser)_3$AlaGlySerAla linker-V_H.4715 and the *Sac*I/*Eco*RI pGOSA.T fragment containing V_L.3418 linked directly to the framework-1 region of V_L.4715 (see Figure 37). The resulting plasmid contains V_H.3418 linked by the $(Gly_4Ser)_3$AlaGlySerAla linker to the 5' end of the framework-1 region of V_H.4715 plus V_L.3418 linked directly to the 5' end of the framework-1 region of V_L.4715, thus:

***pel*B-V_H.3418-linkerA-V_H.4715 + *pel*B-V_L.3418*V_L.4715.**

EXAMPLE 7. Detailed description of the preparation of monocistronic constructs pGOSA.L and pGOSA.Y, and pGOSA.C, pGOSA.X, pGOSA.AC and pGOSA.AD

pGOSA.L

This plasmid is derived from pGOSA.E from which the *Hin*dIII/*Nhe*I fragment containing DNA encoding V_H.4715-$(Gly_4Ser)_3$AlaGlySerAla-V_H.3418 was removed (see Figure 43). The DNA ends of the vector were made blunt-end using Klenow DNA polymerase and ligated. The resulting plasmid pGOSA.L contains V_L.3418 linked by the $(Gly_4Ser)_2Gly_4$Val linker to the 5' end of the framework-1 region of V_L.4715, thus:

***pel*B-V_L.3418-linkerV-V_L.4715.**

pGOSA.Y

This plasmid is derived from pGOSA.T from which the *Hin*dIII/*Nhe*I fragment containing DNA encoding V_H.4715-V_H.3418 was removed (see Figure 44). The DNA ends of the vector were made blunt-end using Klenow DNA polymerase and

ligated. The resulting plasmid pGOSA.Y contains V_L.3418 linked directly to 5′ end of the framework-1 region of V_L.4715, thus:

***pelB*-V_L.3418*V_L.4715.**

The preparation of **pGOSA.C** was given in Example 5 above; it can be indicated with: ***pelB*-V_H.4715-linkerA-V_H.3418.**

pGOSA.X

This plasmid is derived from pGOSA.T from which the *Nhe*I/*Eco*RI fragment containing DNA encoding V_L.3418-V_L.4715 was removed. The DNA ends of the vector were made blunt-end using Klenow DNA polymerase and ligated. The resulting plasmid pGOSA.X (see Figure 45) contains V_H.4715 linked directly to 5′ end of the framework-1 region of V_H.3418, thus: ***pelB*-V_H.4715*V_H.3418.**

pGOSA.AC

This plasmid is derived from pGOSA.Z from which the *Nhe*I/*Eco*RI fragment containing DNA encoding V_L.3418-$(Gly_4Ser)_2Gly_4Val$-V_L.4715 was removed (see Figure 46). The DNA ends of the vector were made blunt-end using Klenow DNA polymerase and ligated. The resulting plasmid pGOSA.AC contains V_H.3418 linked directly to 5′ end of the framework-1 region of V_H.4715, thus:

***pelB*-V_H.3418*V_H.4715.**

pGOSA.AD

This plasmid was obtained by inserting the *Pst*I/*Eco*RI PCR.X fragment containing DNA encoding V_H.3418-$(Gly_4Ser)_3$AlaGly-SerAla-V_H.4715 (see Figure 29) into the Fv.4715-myc vector from which the *Pst*I/*Eco*RI Fv.4715-myc insert was removed (see Figure 47), thus: ***pelB*-V_H.3418-linkerA-V_H.4715.**

These monocistronic constructs can be used to transform the same host with two different plasmids or to transform two

different hosts, so that the two V_H's in series can be produced separately from the two V_L's in series.

Evaluation of the results obtained

Bifunctional binding activity of GOSA double heads

In this specification the construction of a two chain protein complex is described, in which one of the chains consists of two heavy chain V-domains and the other chain consists of the two corresponding light chain V-domains. The variable domains are linked either directly or through a polypeptide linker. In this specification evidence is provided that these type of molecules ("double heads") contain both antigen binding specificities of the Fv's used to generate these multi-functional antibody fragments. Figure 12 shows that GOSA.E can be used to specifically target the enzyme glucose oxidase to several *Streptococcus sanguis* strains, using antibody fragments derived from hybridomas expressing antibodies directed against these antigens. Figure 12 further shows that the fine specificity of the anti-*Streptococcus sanguis* scFv 4715 is preserved in the GOSA.E double head.

Effect of linkers and relative position of V-domains on double head activity

After it was shown that the "cross-over double-head" approach (V_HA-V_HB + V_LB-V_LA) yields active bispecific molecules, the importance of the relative position of the V-domains in these constructs was investigated. Both possible positional orientations (GOSA.E = V_HA-LinkerA-V_HB + V_LB-LinkerV-V_LA and GOSA.J = V_HB-LinkerA-V_HA + V_LB-LinkerV-V_LA) were constructed and tested for bispecific activity, despite the suggestion obtained by molecular modelling that the binding site formed by the second (downstream/C-terminal) V-domains in the configuration V_HB-V_HA + V_LB-V_LA (GOSA.J) was in an unfavourable position for binding to large protein antigens on the surface of cells. Surprisingly however, it was found experimentally that the

downstream binding site is in fact accessible. Although the relative position of the heavy chains and the light chains was found to have an effect on the observed reactivity both tested combinations show bispecific activity with the "cross-over" combination (GOSA.E = V_HA-V_HB + V_LB-V_LA) exhibiting a higher level of reactivity compared to the combination V_HB-V_HA + V_LB-V_LA (= GOSA.J) as demonstrated for A=anti-Strep and B=anti-Gox.

Molecular modelling of the V_HB-V_HA + V_LB-V_LA (= GOSA.J) configuration further suggested that, only when the connecting linkers are kept long enough (to span 30 to 35 Å), the protein chains could fold such that both binding sites are fully accessible.
The "cross-over" configuration: V_HA-V_HB + V_LB-V_LA (GOSA.E) wherein linker length was not critical, was predicted to result in a complex with both binding sites facing in opposite directions, without the restraints suggested for the configuration V_HB-V_HA + V_LB-V_LA (GOSA.J).
Removing the flexible polypeptide linker from the V_HA-V_HB chain only had a minimal effect on the ability of the double head in the "cross-over" configuration (GOSA.V = V_HA*V_HB + V_LB-V_LA) to bind both *S. sanguis* and Glucose oxidase. However, removing the flexible polypeptide linker from the V_HB-V_HA chain from the molecule in the V_HB-V_HA + V_LB-V_LA configuration (GOSA.Z = V_HB*V_HA + V_LB-V_LA) resulted in a dramatic reduction of its ability to bind both *S. sanguis* and Glucose oxidase.
In contrast with the double head in the "cross-over" configuration without the flexible polypeptide linker between the two heavy chain domains (GOSA.V), where molecular modelling predicted the resulting molecule to be active, removal of the flexible linker from the V_LB-V_LA chain could not be modelled such that both binding sites were fully accessible. ELISA results confirm that the double head in the V_HB-V_HA + V_LB-V_LA configuration without a linker between the two light chain domains (GOSA.AB)

exhibits only minimal *S. sanguis* and glucose oxidase binding activity. Surprisingly, deletion of the flexible linker from the V_LB-V_LA chain from the double head in the "cross-over" configuration (GOSA.S) only had a small effect on the bispecific activity of the resulting molecule.
As expected from the molecular modelling results from the double heads without a flexible linker between the two light chain domains, removal of both the flexible polypeptide linkers from the double head molecules, could not be modelled such that both binding sites were fully accessible. In agreement with the ELISA results obtained with the GOSA.AB construct, the double head in the V_HB-V_HA + V_LB-V_LA configuration without any linkers (GOSA.AA) only exhibits minimal if any *S. sanguis* and glucose oxidase binding activity. Surprisingly, the double head in the "cross-over" configuration without any linkers (GOSA.T = V_HA*V_HB + V_LB*V_LA) still exhibited 25-50% of *S. sanguis* and glucose oxidase bispecific binding activity when compared to the activity of the double head in the "cross-over" configuration with two linkers (GOSA.E).
Thus the conclusion of this work is that modelling can give some indications, but that the computer programmes cannot predict what is possible and what not. Several deviations from the modelling expectations were found. With a paraphrase on an old saying: theories are nice but the experiment is the ultimate proof.

Sensitivity of GOSA double heads
Using an ELISA format it was shown that the sensitivity of the GOSA.E double head is as least as a sensitive as an IgG-glucose oxidase conjugate, as determined by the lowest concentration of *Streptococcus sanguis* antigen immobilised on a solid phase that is still detectable.

GOSA double heads are produced as dimers
FPLC analysis of partially affinity-purified GOSA.E, GOSA.V, GOSA.S and GOSA.T samples usually gave only one

GOSA double head activity peak as determined by ELISA (Figures 16-19). The position of this peak in the elution pattern indicated that the molecular weight of the GOSA double head is 40-50 kD. Since this molecular weight corresponds to the expected molecular weight of the V_H2 + V_L2 double head dimer, it was concluded that GOSA.E, GOSA.V, GOSA.S and GOSA.T are primarily produced as dimeric molecules. Occasionally an activity peak with an apparent molecular weight of ≈200 kD was observed (Figure 16). The presence of glucose oxidase activity in these fractions indicate that these fractions contain GOSA double head complexed with glucose oxidase.

In vitro assembly of GOSA double heads

It was shown that bifunctionally active dimeric GOSA molecules together in one cell can be produced by translation from one dicistronic messenger (GOSA.E, GOSA.S, GOSA.T, GOSA.V, GOSA.J, GOSA.AB, GOSA.AA and GOSA.Z). In addition high levels of *S. sanguis* and glucose oxidase bispecific binding activity is formed when supernatants of cultures producing the separate GOSA subunits are mixed (see Example 7). The effects of linkers and the relative position of the individual V_H-domains on the *S. sanguis* and glucose oxidase bispecific binding activity observed in these mixing experiments are comparable to the dicistronic constructs.

The constructs described above are summarised in Table 2 below.

Table 2A describes intermediate constructs that were not further tested.

Table 2B describes the dicistronic constructs.

Table 2C describes the monocistronic constructs.

(LiA) stands for the V_H-V_H linker $(Gly_4Ser)_3$AlaGlySerAla (= linkerA)

(LiV) stands for the V_L-V_L linker $(Gly_4Ser)_2Gly_4$Val (= linkerV)

(*) indicates that the two heavy chain domains or the two light chain domains are fused together without a connecting linker.

Table 2.

Table 2A

GOSA.A :	V_H.4715-LiA-(*Sfi*I)-V_L.4715-myc
GOSA.B :	V_H.3418-LiV-V_L.3418-(*Sal*I/*Eco*RI)
GOSA.D :	V_H.3418 + V_L.3418-LiV-V_L.4715
GOSA.G :	V_H.3418-LiA-V_H.3418 + V_L.3418-LiV-V_L.4715

Table 2B

GOSA.E :	V_H.4715-LiA-V_H.3418 + V_L.3418-LiV-V_L.4715
GOSA.S :	V_H.4715-LiA-V_H.3418 + V_L.3418*V_L.4715
GOSA.T :	V_H.4715*V_H.3418 + V_L.3418*V_L.4715
GOSA.V :	V_H.4715*V_H.3418 + V_L.3418-LiV-V_L.4715
GOSA.J :	V_H.3418-LiA-V_H.4715 + V_L.3418-LiV-V_L.4715
GOSA.AB:	V_H.3418-LiA-V_H.4715 + V_L.3418*V_L.4715
GOSA.AA:	V_H.3418*V_H.4715 + V_L.3418*V_L.4715
GOSA.Z :	V_H.3418*V_H.4715 + V_L.3418-LiV-V_L.4715

Table 2C

GOSA.L :	V_L.3418-LiV-V_L.4715
GOSA.Y :	V_L.3418*V_L.4715
GOSA.AD:	V_H.3418-LiA-V_H.4715
GOSA.AC:	V_H.3418*V_H.4715
GOSA.C :	V_H.4715-LiA-V_H.3418
GOSA.X :	V_H.4715*V_H.3418

REFERENCES

Patent literature

- EP-0281604; GENEX/ENZON; SINGLE POLYPEPTIDE CHAIN BINDING MOLECULES; priority date 02.09.86.
- WO 93/11161; ENZON, INC; MULTIVALENT ANTIGEN-BINDING PROTEINS; priority date 25.11.91.
- WO 94/09131; SCOTGEN LIMITED; RECOMBINANT SPECIFIC BINDING PROTEIN; priority date 15.10.92).
- WO 94/13804 (CAMBRIDGE ANTIBODY TECHNOLOGY / MEDICAL RESEARCH COUNCIL; MULTIVALENT AND MULTISPECIFIC BINDING PROTEINS, THEIR MANUFACTURE AND USE; first priority date 04.12.92
- WO 94/13806; THE DOW CHEMICAL COMPANY; MULTIVALENT SINGLE CHAIN ANTIBODIES; priority date 11.12.92
- WO 94/25591; UNILEVER N.V., UNILEVER PLC, Hamers, R., Hamers-Casterman, C., & Muyldermans, S; PRODUCTION OF ANTIBODIES OR (FUNCTIONALIZED) FRAGMENTS THEREOF DERIVED FROM HEAVY CHAIN IMMUNOGLOBULINS OF *CAMELIDAE*; first priority date 29.04.93.

Non-patent literature

- Anthony, J., Near, R., Wong, S.L., Iida, E., Ernst, E., Wittekind, M., Haber, E. and Ng, S-C; **Molec. Immunol.** 29 (1992) 1237-1247, PRODUCTION OF STABLE ANTI-DIGOXIN Fv IN *ESCHERICHIA COLI*.
- Berry, M.J. and Davies, J.; **J. Chromatography** 597 (1992) 239-245; Use of antibody fragments in immunoaffinity chromatography: comparison of Fv fragments, V_H fragments and paralog peptides.
- Better, M., Chang, C.P., Robinson, R.R. and Horwitz A.H.; **Science** 240 (1988) 1041-1043; *Escherichia coli* Secretion of an Active Chimeric Antibody Fragment.
- Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S. and Whitlow, M.; **Science** 242 (1988) 423-426; Single-Chain Antigen-Binding Proteins.

- Carter, P., Kelley, R.F., Rodrigues, M.L., Snedecor, B., Covarrubias, M., Velligan, M.D., Wong, W.L.T., Rowland, A.M., Kotts, C.E., Carver, M.E., Yang, M., Bourell, J.H., Shepard, H.M. and Henner, D.; **BIO/TECHNOLOGY** 10 (1992) 163-167; High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment.

- Cumber, A.J., Ward, E.S., Winter, G., Parnell, G.D. and Wawrzynczak. E.J.; **J. Immunol.** 149 (1992) 120-126; COMPARATIVE STABILITIES IN VITRO AND IN VIVO OF A RECOMBINANT MOUSE ANTIBODY FvCys FRAGMENT AND A bisFvCys CONJUGATE.

- Firek, S., Draper, J., Owen. M.R.L., Gandecha, A., Cockburn, B., and Whitelam, G.C.; **Plant Mol. Biol.** 23 (1993) 861-870; Secretion of a functional single-chain Fv protein in transgenic tobacco plants and cell suspension cultures.

- Givol, D.; **Molec. Immunol.** 28 (1991) 1379; THE MINIMAL ANTIGEN BINDING FRAGMENT OF ANTIBODIES - Fv FRAGMENT.

- Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Bajyana Songa, E., Bendahman, N. and Hamers, R.; **Nature** 363 (1993) 446-448; Naturally occurring antibodies devoid of light chains.

- Hiatt, A., Cafferkey, R. and Bowdish, K.; **Nature** 342 (1989) 76-78; Production of antibodies in transgenic plants.

- Holliger, P., Prospero, T., and Winter, G.; **Proc. Natl. Acad. Sci. USA** 90 (1993) 6444-6448; "Diabodies": Small bivalent and bispecific antibody fragments.

- Horwitz, A.H., Chang, C.P., Better, M., Hellstrom, K.E. and Robinson, R.R.; **Proc. Natl. Acad. Sci. USA** 85 (1988) 8678-8682; Secretion of functional antibody and Fab fragment from yeast cells.

- Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M-S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R. and Oppermann, H.; **Proc. Natl. Acad. Sci. USA** 85 (1988) 5879-5883; Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*.

- Kostelny, S.A., Cole, M.S. and Tso, J.Y.; **J. Immunol.** 148 (1992) 1547-1553; FORMATION OF A BISPECIFIC ANTIBODY BY THE USE OF LEUCINE ZIPPERS.

- Mallender, W.D., and Voss, Jr., E.W.; **J. Biol. Chem.** 269 (1994) 199-206; Construction, Expression, and Activity of a Bivalent Bispecific Single-chain Antibody.

- Milstein, C. and Cuello, A.C.; **Nature,** 305 (1983) 537-540; Hybrid hybridomas and their use in immunohistochemistry.

- Nyyssönen, E., Penttillä, M., Harkki, A., Saloheimo, A., Knowles, J.K.C. and Keränen, S.; **BIO/TECHNOLOGY Bio/tech.** 11 (1993) 591-595; Efficient Production of Antibody Fragments in Filamentous Fungus *Trichoderma reesei.*

- Owen, M., Gandecha, A., Cockburn, B. and Whitelam G.; **BIO/TECHNOLOGY** 10 (1992) 790-794; Synthesis of a functional anti-phytochrome single-chain Fv protein in transgenic tobacco.

- Pack, P. and Plückthun, A.; **Biochemistry** 31 (1992) 1579-1584; Miniantibodies: Use of Amphiphatic Helices to Produce Functional, Flexibly Linked Dimeric Fv Fragments with High Avidity in *Escherichia coli.*

- Skerra, A. and Plückthun, A.; **Science,** 240 (1988) 1038-1041; Assembly of a Functional Immunoglobulin Fv Fragment in *Escherichia coli.*

- Taub, R., Gould, R.J., Ciccarone, T.M., Hoxie, J., Friedman, P.A., Shattil, S.J. and Garsky, V.M.; **J. Biol. Chem.** 264 (1989) 259-265; A Monoclonal Antibody against the Platelet Fibrinogen Receptor Contains a Sequence That Mimics a Receptor Recognition Domain in Fibrinogen.

- Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T. and Winter G.; **Nature** 341 (1989) 544-546; Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli.*

- Williams, W,V., Moss, D.A., Kieber-Emmons, T., Cohen, J.A., Myers, J.N., Weiner, D.B. and Greene, M.I.; **Proc. Natl. Acad. Sci. USA.** 86 (1989) 5537-5541; Development of biologically active peptides based on antibody structure.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Unilever PLC
(B) STREET: Unilever House, Blackfriars
(C) CITY: London
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): EC4P 4BQ (GB)

(A) NAME: Unilever N.V.
(B) STREET: Weena 455
(C) CITY: Rotterdam
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): NL-3013 AL

(A) NAME: Paul James DAVIS
(B) STREET: The Hawthorns, Pavenham Road
(C) CITY: Felmersham (Bedfordshire)
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): MK43 7EX (GB)

(A) NAME: Cornelis Paul Erik van der LOGT
(B) STREET: 1 Bluebell Rise (Peverel Manor Estate)
(C) CITY: Rushden (Northamptonshire)
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): NN10 OTU (GB)

(A) NAME: Martine Elisa VERHOEIJEN
(B) STREET: 1 Tintagel Close (Manor Farm Estate)
(C) CITY: Rushden (Northamptonshire)
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): NN10 ONP (GB)

(A) NAME: Steve Wilson
(B) STREET: 3 Aldenham Close (Goldington)
(C) CITY: Bedford,
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): MK41 OFQ (GB)

(ii) TITLE OF INVENTION: A bifunctional or bivalent antibody fragment analogue

(iii) NUMBER OF SEQUENCES: 31

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: EP 95307332.7

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 737 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

(vii) IMMEDIATE SOURCE:
(B) CLONE: EcoRI-HindIII insert of pUR4124

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:11..730
(D) OTHER INFORMATION:/product= "VLlys-GS-VHlys"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:11..334
(D) OTHER INFORMATION:/product= "VLlys"

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:335..379
(D) OTHER INFORMATION:/product= "(Gly4Ser)3 linker"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:380..727
(D) OTHER INFORMATION:/product= "VHlys"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
GAATTCGGCC GAC ATC GAG CTC ACC CAG TCT CCA GCC TCC CTT TCT GCG     49
           Asp Ile Glu Leu Thr Gln Ser Pro Ala Ser Leu Ser Ala
             1               5                  10

TCT GTG GGA GAA ACT GTC ACC ATC ACA TGT CGA GCA AGT GGG AAT ATT    97
Ser Val Gly Glu Thr Val Thr Ile Thr Cys Arg Ala Ser Gly Asn Ile
     15                  20                  25

CAC AAT TAT TTA GCA TGG TAT CAG CAG AAA CAG GGA AAA TCT CCT CAG   145
His Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro Gln
 30                  35                  40                  45

CTC CTG GTC TAT TAT ACA ACA ACC TTA GCA GAT GGT GTG CCA TCA AGG   193
Leu Leu Val Tyr Tyr Thr Thr Thr Leu Ala Asp Gly Val Pro Ser Arg
                 50                  55                  60

TTC AGT GGC AGT GGA TCA GGA ACA CAA TAT TCT CTC AAG ATC AAC AGC   241
Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys Ile Asn Ser
             65                  70                  75

CTG CAA CCT GAA GAT TTT GGG AGT TAT TAC TGT CAA CAT TTT TGG AGT   289
Leu Gln Pro Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His Phe Trp Ser
         80                  85                  90

ACT CCT CGG ACG TTC GGT GGA GGG ACC AAG CTC GAG ATC AAA CGG GGT   337
Thr Pro Arg Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly
     95                 100                 105

GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC GGA TCG CAG GTG   385
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gln Val
110                 115                 120                 125

CAG CTG CAG GAG TCA GGA CCT GGC CTG GTG GCG CCC TCA CAG AGC CTG   433
Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu
                130                 135                 140

TCC ATC ACA TGC ACC GTC TCA GGG TTC TCA TTA ACC GGC TAT GGT GTA   481
Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val
            145                 150                 155
```

```
AAC TGG GTT CGC CAG CCT CCA GGA AAG GGT CTG GAG TGG CTG GGA ATG        529
Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met
        160                 165                 170

ATT TGG GGT GAT GGA AAC ACA GAC TAT AAT TCA GCT CTC AAA TCC AGA        577
Ile Trp Gly Asp Gly Asn Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg
    175                 180                 185

CTG AGC ATC AGC AAG GAC AAC TCC AAG AGC CAA GTT TTC TTA AAA ATG        625
Leu Ser Ile Ser Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met
190                 195                 200                 205

AAC AGT CTG CAC ACT GAT GAC ACA GCC AGG TAC TAC TGT GCC AGA GAG        673
Asn Ser Leu His Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Glu
                210                 215                 220

AGA GAT TAT AGG CTT GAC TAC TGG GGC CAA GGG ACC ACG GTC ACC GTC        721
Arg Asp Tyr Arg Leu Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val
            225                 230                 235

TCC TCA TGA TAAGCTT                                                    737
Ser Ser  *
        240
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Asp Ile Glu Leu Thr Gln Ser Pro Ala Ser Leu Ser Ala Ser Val Gly
1               5                   10                  15

Glu Thr Val Thr Ile Thr Cys Arg Ala Ser Gly Asn Ile His Asn Tyr
            20                  25                  30

Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro Gln Leu Leu Val
        35                  40                  45

Tyr Tyr Thr Thr Thr Leu Ala Asp Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys Ile Asn Ser Leu Gln Pro
65                  70                  75                  80

Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His Phe Trp Ser Thr Pro Arg
                85                  90                  95

Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly Gly Gly Gly
            100                 105                 110

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Gln
        115                 120                 125

Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu Ser Ile Thr
    130                 135                 140

Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val Asn Trp Val
145                 150                 155                 160
```

```
Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met Ile Trp Gly
                165                 170                 175

Asp Gly Asn Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg Leu Ser Ile
            180                 185                 190

Ser Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met Asn Ser Leu
        195                 200                 205

His Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Glu Arg Asp Tyr
    210                 215                 220

Arg Leu Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser  *
225                 230                 235                     240
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 920 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

(vii) IMMEDIATE SOURCE:
(B) CLONE: HindIII-EcoRI insert Fv.3418

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:36..443
(D) OTHER INFORMATION:/product= "pelB-VH3418"

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION:36..101
(D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:102..440
(D) OTHER INFORMATION:/product= "VH3418"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:495..884
(D) OTHER INFORMATION:/product= "pelB-VL4318"

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION:495..560
(D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:561..881
(D) OTHER INFORMATION:/product= "VL3418"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
AAGCTTGCAA ATTCTATTTC AAGGAGACAG TCATA ATG AAA TAC CTA TTG CCT      53
                                       Met Lys Tyr Leu Leu Pro
                                       -22     -20
```

```
ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG GCC    101
Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met Ala
    -15                 -10                     -5

CAG GTG CAG CTG CAG CAG TCA GGA CCT GAG CTG GTA AAG CCT GGG GCT    149
Gln Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala
 1               5                  10                  15

TCA GTG AAG ATG TCC TGC AAG GCT TCT GGA TAC ACA TTC ACT AGC TAT    197
Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
            20                  25                  30

GTT ATG CAC TGG GTG AAA CAG AAG CCT GGG CAG GGC CTT GAG TGG ATT    245
Val Met His Trp Val Lys Gln Lys Pro Gly Gln Gly Leu Glu Trp Ile
        35                  40                  45

GGA TAT ATT TAT CCT TAC AAT GAT GGT ACT AAG TAC AAT GAG AAG TTC    293
Gly Tyr Ile Tyr Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys Phe
    50                  55                  60

AAA GGC AAG GCC ACA CTG ACT TCA GAC AAA TCC TCC AGC ACA GCC TAC    341
Lys Gly Lys Ala Thr Leu Thr Ser Asp Lys Ser Ser Ser Thr Ala Tyr
 65                 70                  75                  80

ATG GAG CTC AGC AGC CTG ACC TCT GAG GAC TCT GCG GTC TAT TAC TGT    389
Met Glu Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
                85                  90                  95

TCA AGA CGC TTT GAC TAC TGG GGC CAA GGG ACC ACG GTC ACC GTC TCC    437
Ser Arg Arg Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser
            100                 105                 110

TCA TAA TAAGAGCTAT GGGAGCTTGC ATGCAAATTC TATTTCAAGG AGACAGTCAT    493
Ser  *

A ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC      539
  Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu
  -22     -20                 -15                 -10

GCT GCC CAA CCA GCG ATG GCC GAC ATC GAG CTC ACC CAG TCT CCA TCT    587
Ala Ala Gln Pro Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Ser
        -5                   1               5

TCC ATG TAT GCA TCT CTA GGA GAG AGA ATC ACT ATC ACT TGC AAG GCG    635
Ser Met Tyr Ala Ser Leu Gly Glu Arg Ile Thr Ile Thr Cys Lys Ala
10                  15                  20                  25

AGT CAG GAC ATT AAT ACC TAT TTA ACC TGG TTC CAG CAG AAA CCA GGG    683
Ser Gln Asp Ile Asn Thr Tyr Leu Thr Trp Phe Gln Gln Lys Pro Gly
                30                  35                  40

AAA TCT CCC AAG ACC CTG ATC TAT CGT GCA AAC AGA TTG CTA GAT GGG    731
Lys Ser Pro Lys Thr Leu Ile Tyr Arg Ala Asn Arg Leu Leu Asp Gly
            45                  50                  55

GTC CCA TCA AGG TTC AGT GGC AGT GGA TCT GGG CAA GAT TAT TCT CTC    779
Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu
        60                  65                  70

ACC ATC AGC AGC CTG GAC TAT GAA GAT ATG GGA ATT TAT TAT TGT CTA    827
Thr Ile Ser Ser Leu Asp Tyr Glu Asp Met Gly Ile Tyr Tyr Cys Leu
    75                  80                  85

CAA TAT GAT GAG TTG TAC ACG TTC GGA GGG GGG ACC AAG CTC GAG ATC    875
Gln Tyr Asp Glu Leu Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
90                  95                  100                 105
```

```
AAA CGG TAA TAATGATCAA ACGGTATAAG GATCCAGCTC GAATTC                  920
Lys Arg  *
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 136 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22     -20                 -15                 -10

Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Pro Glu
    -5                  1               5                   10

Leu Val Lys Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly
                15                  20                  25

Tyr Thr Phe Thr Ser Tyr Val Met His Trp Val Lys Gln Lys Pro Gly
            30                  35                  40

Gln Gly Leu Glu Trp Ile Gly Tyr Ile Tyr Pro Tyr Asn Asp Gly Thr
        45                  50                  55

Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Ser Asp Lys
    60                  65                  70

Ser Ser Ser Thr Ala Tyr Met Glu Leu Ser Ser Leu Thr Ser Glu Asp
75                  80                  85                  90

Ser Ala Val Tyr Tyr Cys Ser Arg Arg Phe Asp Tyr Trp Gly Gln Gly
                95                  100                 105

Thr Thr Val Thr Val Ser Ser  *
            110
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 130 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22     -20                 -15                 -10

Ala Gln Pro Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Ser Ser
    -5                  1               5                   10

Met Tyr Ala Ser Leu Gly Glu Arg Ile Thr Ile Thr Cys Lys Ala Ser
                15                  20                  25

Gln Asp Ile Asn Thr Tyr Leu Thr Trp Phe Gln Gln Lys Pro Gly Lys
            30                  35                  40
```

```
Ser Pro Lys Thr Leu Ile Tyr Arg Ala Asn Arg Leu Leu Asp Gly Val
        45                  50                  55

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu Thr
    60                  65                  70

Ile Ser Ser Leu Asp Tyr Glu Asp Met Gly Ile Tyr Tyr Cys Leu Gln
75                  80                  85                  90

Tyr Asp Glu Leu Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
                95                  100                 105

Arg  *
```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

(vii) IMMEDIATE SOURCE:
(B) CLONE: HindIII-EcoRI insert of Fv.4715-myc

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:40..468
(D) OTHER INFORMATION:/product= "pelB-VH4715"

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION:40..105
(D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:106..465
(D) OTHER INFORMATION:/product= "VH4715"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:520..963
(D) OTHER INFORMATION:/product= "pelB-VL4715-myc"

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION:520..585
(D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:586..927
(D) OTHER INFORMATION:/product= "VL4715"

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:928..960
(D) OTHER INFORMATION:/product= "myc-tag"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATA ATG AAA TAC CTA TTG        54
                                          Met Lys Tyr Leu Leu
                                          -22     -20

CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG      102
Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met
    -15                 -10                 -5

GCC CAG GTG CAG CTG CAG GAG TCA GGG GGA GAC TTA GTG AAG CCT GGA      150
Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp Leu Val Lys Pro Gly
    1               5                   10                  15

GGG TCC CTG ACA CTC TCC TGT GCA ACC TCT GGA TTC ACT TTC AGT AGT      198
Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser Ser
                20                  25                  30

TAT GCC TTT TCT TGG GTC CGC CAG ACC TCA GAC AAG AGT CTG GAG TGG      246
Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp Lys Ser Leu Glu Trp
            35                  40                  45

GTC GCA ACC ATC AGT AGT ACT GAT ACT TAT ACC TAT TAT TCA GAC AAT      294
Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr Tyr Tyr Ser Asp Asn
        50                  55                  60

GTG AAG GGC CGC TTC ACC ATC TCC AGA GAC AAT GGC AAG AAC ACC CTG      342
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Gly Lys Asn Thr Leu
    65                  70                  75

TAC CTG CAA ATG AGC AGT CTG AAG TCT GAG GAC ACA GCC GTG TAT TAC      390
Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp Thr Ala Val Tyr Tyr
80                  85                  90                  95

TGT GCA AGA CAT GGG TAC TAT GGT AAA GGC TAT TTT GAC TAC TGG GGC      438
Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr Phe Asp Tyr Trp Gly
                100                 105                 110

CAA GGG ACC ACG GTC ACC GTC TCC TCA TAA TAAGAGCTAT GGGAGCTTGC        488
Gln Gly Thr Thr Val Thr Val Ser Ser  *
            115                 120

ATGCAAATTC TATTTCAAGG AGACAGTCAT A ATG AAA TAC CTA TTG CCT ACG       540
                                   Met Lys Tyr Leu Leu Pro Thr
                                   -22     -20

GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG GCC GAC      588
Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Asp
-15                 -10                 -5                  1

ATC GAG CTC ACT CAG TCT CCA TTC TCC CTG ACT GTG ACA GCA GGA GAG      636
Ile Glu Leu Thr Gln Ser Pro Phe Ser Leu Thr Val Thr Ala Gly Glu
            5                   10                  15

AAG GTC ACT ATG AAT TGC AAG TCC GGT CAG AGT CTG TTA AAC AGT GTA      684
Lys Val Thr Met Asn Cys Lys Ser Gly Gln Ser Leu Leu Asn Ser Val
        20                  25                  30

AAT CAG AGG AAC TAC TTG ACC TGG TAC CAG CAG AAG CCA GGG CAG CCT      732
Asn Gln Arg Asn Tyr Leu Thr Trp Tyr Gln Gln Lys Pro Gly Gln Pro
    35                  40                  45

CCT AAA CTG TTG ATC TAC TGG GCA TCC ACT AGG GAA TCT GGA GTC CCT      780
Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro
50                  55                  60                  65
```

```
GAT CGC TTC ACA GCC AGT GGA TCT GGA ACA GAT TTC ACT CTC ACC ATC      828
Asp Arg Phe Thr Ala Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile
                 70                      75                  80

AGC AGT GTG CAG GCT GAA GAC CTG GCA GTT TAT TAC TGT CAG AAT GAT      876
Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn Asp
             85                      90                  95

TAT ACT TAT CCG TTC ACG TTC GGA GGG GGG ACC AAG CTC GAG ATC AAA      924
Tyr Thr Tyr Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
        100                 105                 110

CGG GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT TAA TAAGATCAAA      973
Arg Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn  *
    115                 120                 125

CGGTAATAAG GATCCAGCTC GAATTC                                        999
```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 143 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22     -20                 -15                 -10

Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp
     -5                   1               5                  10

Leu Val Lys Pro Gly Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly
                 15                  20                  25

Phe Thr Phe Ser Ser Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp
             30                  35                  40

Lys Ser Leu Glu Trp Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr
         45                  50                  55

Tyr Tyr Ser Asp Asn Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
     60                  65                  70

Gly Lys Asn Thr Leu Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp
 75                  80                  85                  90

Thr Ala Val Tyr Tyr Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr
                 95                 100                 105

Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser  *
            110                 115                 120
```

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 148 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22     -20                 -15                 -10

Ala Gln Pro Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Phe Ser
    -5                   1               5                  10

Leu Thr Val Thr Ala Gly Glu Lys Val Thr Met Asn Cys Lys Ser Gly
                15                  20                  25

Gln Ser Leu Leu Asn Ser Val Asn Gln Arg Asn Tyr Leu Thr Trp Tyr
            30                  35                  40

Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser
        45                  50                  55

Thr Arg Glu Ser Gly Val Pro Asp Arg Phe Thr Ala Ser Gly Ser Gly
    60                  65                  70

Thr Asp Phe Thr Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala
75                  80                  85                  90

Val Tyr Tyr Cys Gln Asn Asp Tyr Thr Tyr Pro Phe Thr Phe Gly Gly
                95                  100                 105

Gly Thr Lys Leu Glu Ile Lys Arg Glu Gln Lys Leu Ile Ser Glu Glu
            110                 115                 120

Asp Leu Asn  *
        125
```

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 924 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

(vii) IMMEDIATE SOURCE:
(B) CLONE: HindIII-EcoRI insert of scFv.4715-myc

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION:40..105
(D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:106..465
(D) OTHER INFORMATION:/product= "VH4715"

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:466..510
(D) OTHER INFORMATION:/product= "(Gly4Ser)3-linker"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:511..852
(D) OTHER INFORMATION:/product= "VL4715"

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:853..885
(D) OTHER INFORMATION:/product= "myc-tag"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:40..888
(D) OTHER INFORMATION:/product= "pelB-VH4715-(Gly4Ser)3-VL4715-myc"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATA ATG AAA TAC CTA TTG      54
                                          Met Lys Tyr Leu Leu
                                          -22     -20

CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG    102
Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met
        -15                 -10                 -5

GCC CAG GTG CAG CTG CAG GAG TCA GGG GGA GAC TTA GTG AAG CCT GGA    150
Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp Leu Val Lys Pro Gly
    1               5                   10                  15

GGG TCC CTG ACA CTC TCC TGT GCA ACC TCT GGA TTC ACT TTC AGT AGT    198
Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser Ser
                20                  25                  30

TAT GCC TTT TCT TGG GTC CGC CAG ACC TCA GAC AAG AGT CTG GAG TGG    246
Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp Lys Ser Leu Glu Trp
            35                  40                  45

GTC GCA ACC ATC AGT AGT ACT GAT ACT TAT ACC TAT TAT TCA GAC AAT    294
Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr Tyr Tyr Ser Asp Asn
        50                  55                  60

GTG AAG GGG CGC TTC ACC ATC TCC AGA GAC AAT GGC AAG AAC ACC CTG    342
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Gly Lys Asn Thr Leu
    65                  70                  75

TAC CTG CAA ATG AGC AGT CTG AAG TCT GAG GAC ACA GCC GTG TAT TAC    390
Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp Thr Ala Val Tyr Tyr
80                  85                  90                  95

TGT GCA AGA CAT GGG TAC TAT GGT AAA GGC TAT TTT GAC TAC TGG GGC    438
Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr Phe Asp Tyr Trp Gly
                100                 105                 110

CAA GGG ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA    486
Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly
            115                 120                 125

GGT GGC TCT GGC GGT GGC GGA TCG GAC ATC GAG CTC ACT CAG TCT CCA    534
Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Glu Leu Thr Gln Ser Pro
        130                 135                 140

TTC TCC CTG ACT GTG ACA GCA GGA GAG AAG GTC ACT ATG AAT TGC AAG    582
Phe Ser Leu Thr Val Thr Ala Gly Glu Lys Val Thr Met Asn Cys Lys
    145                 150                 155

TCC GGT CAG AGT CTG TTA AAC AGT GTA AAT CAG AGG AAC TAC TTG ACC    630
Ser Gly Gln Ser Leu Leu Asn Ser Val Asn Gln Arg Asn Tyr Leu Thr
160                 165                 170                 175
```

TGG TAC CAG CAG AAG CCA GGG CAG CCT CCT AAA CTG TTG ATC TAC TGG 678
Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Trp
180 185 190

GCA TCC ACT AGG GAA TCT GGA GTC CCT GAT CGC TTC ACA GCC AGT GGA 726
Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg Phe Thr Ala Ser Gly
195 200 205

TCT GGA ACA GAT TTC ACT CTC ACC ATC AGC AGT GTG CAG GCT GAA GAC 774
Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Val Gln Ala Glu Asp
210 215 220

CTG GCA GTT TAT TAC TGT CAG AAT GAT TAT ACT TAT CCG TTC ACG TTC 822
Leu Ala Val Tyr Tyr Cys Gln Asn Asp Tyr Thr Tyr Pro Phe Thr Phe
225 230 235

GGA GGG GGG ACC AAG CTC GAG ATC AAA CGG GAA CAA AAA CTC ATC TCA 870
Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Glu Gln Lys Leu Ile Ser
240 245 250 255

GAA GAG GAT CTG AAT TAA TAAGATCAAA CGGTAATAAG GATCCAGCTC GAATTC 924
Glu Glu Asp Leu Asn *
260

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 283 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22 -20 -15 -10

Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp
-5 1 5 10

Leu Val Lys Pro Gly Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly
15 20 25

Phe Thr Phe Ser Ser Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp
30 35 40

Lys Ser Leu Glu Trp Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr
45 50 55

Tyr Tyr Ser Asp Asn Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
60 65 70

Gly Lys Asn Thr Leu Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp
75 80 85 90

Thr Ala Val Tyr Tyr Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr
95 100 105

Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly
110 115 120

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Glu
125 130 135

Leu Thr Gln Ser Pro Phe Ser Leu Thr Val Thr Ala Gly Glu Lys Val
140 145 150

Thr Met Asn Cys Lys Ser Gly Gln Ser Leu Leu Asn Ser Val Asn Gln
155   160   165   170

Arg Asn Tyr Leu Thr Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys
                175   180   185

Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg
            190   195   200

Phe Thr Ala Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
        205   210   215

Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn Asp Tyr Thr
    220   225   230

Tyr Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Glu
235   240   245   250

Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn *
                255   260

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1706 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

(vii) IMMEDIATE SOURCE:
(B) CLONE: HindIII-EcoRI insert of pGOSA.E

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:40..864
(D) OTHER INFORMATION:/product= "pelB-VH4715-LiA-VH3418"

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION:40..105
(D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:106..465
(D) OTHER INFORMATION:/product= "VH4715"

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:466..522
(D) OTHER INFORMATION:/product= "linkerA
(Gly4Ser)3AlaGlySerAla"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:523..861
(D) OTHER INFORMATION:/product= "VH3418"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:913..1689
(D) OTHER INFORMATION:/product= "pelB-VL3418-LiV-VL4715"

(ix) FEATURE:
 (A) NAME/KEY: sig_peptide
 (B) LOCATION:913..978
 (D) OTHER INFORMATION:/product= "pectate lyase"

(ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION:979..1299
 (D) OTHER INFORMATION:/product= "VL3418"

(ix) FEATURE:
 (A) NAME/KEY: misc_RNA
 (B) LOCATION:1300..1344
 (D) OTHER INFORMATION:/product= "linker V
 (Gly4Ser)2Gly4Val"

(ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION:1345..1686
 (D) OTHER INFORMATION:/product= "VL4715"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```
AAGCTTGCAT GGAAATTCTA TTTCAAGGAG ACAGTCATA ATG AAA TAC CTA TTG       54
                                           Met Lys Tyr Leu Leu
                                           -22     -20

CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG     102
Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met
        -15                 -10                 -5

GCC CAG GTG CAG CTG CAG GAG TCA GGG GGA GAC TTA GTG AAG CCT GGA     150
Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp Leu Val Lys Pro Gly
      1               5                  10                  15

GGG TCC CTG ACA CTC TCC TGT GCA ACC TCT GGA TTC ACT TTC AGT AGT     198
Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser Ser
                 20                  25                  30

TAT GCC TTT TCT TGG GTC CGC CAG ACC TCA GAC AAG AGT CTG GAG TGG     246
Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp Lys Ser Leu Glu Trp
             35                  40                  45

GTC GCA ACC ATC AGT AGT ACT GAT ACT TAT ACC TAT TAT TCA GAC AAT     294
Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr Tyr Tyr Ser Asp Asn
         50                  55                  60

GTG AAG GGG CGC TTC ACC ATC TCC AGA GAC AAT GGC AAG AAC ACC CTG     342
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Gly Lys Asn Thr Leu
     65                  70                  75

TAC CTG CAA ATG AGC AGT CTG AAG TCT GAG GAC ACA GCC GTG TAT TAC     390
Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp Thr Ala Val Tyr Tyr
 80                  85                  90                  95

TGT GCA AGA CAT GGG TAC TAT GGT AAA GGC TAT TTT GAC TAC TGG GGC     438
Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr Phe Asp Tyr Trp Gly
                100                 105                 110

CAA GGG ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA     486
Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly
            115                 120                 125

GGT GGC TCT GGC GGT GGC GGA TCG GCC GGT TCG GCC CAG GTC CAG CTG     534
Gly Gly Ser Gly Gly Gly Gly Ser Ala Gly Ser Ala Gln Val Gln Leu
        130                 135                 140
```

```
CAA CAG TCA GGA CCT GAG CTG GTA AAG CCT GGG GCT TCA GTG AAG ATG      582
Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala Ser Val Lys Met
    145                 150                 155

TCC TGC AAG GCT TCT GGA TAC ACA TTC ACT AGC TAT GTT ATG CAC TGG      630
Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Val Met His Trp
160                 165                 170                 175

GTG AAA CAG AAG CCT GGG CAG GGC CTT GAG TGG ATT GGA TAT ATT TAT      678
Val Lys Gln Lys Pro Gly Gln Gly Leu Glu Trp Ile Gly Tyr Ile Tyr
                180                 185                 190

CCT TAC AAT GAT GGT ACT AAG TAC AAT GAG AAG TTC AAA GGC AAG GCC      726
Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala
            195                 200                 205

ACA CTG ACT TCA GAC AAA TCC TCC AGC ACA GCC TAC ATG GAG CTC AGC      774
Thr Leu Thr Ser Asp Lys Ser Ser Ser Thr Ala Tyr Met Glu Leu Ser
        210                 215                 220

AGC CTG ACC TCT GAG GAC TCT GCG GTC TAT TAC TGT TCA AGA CGC TTT      822
Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ser Arg Arg Phe
    225                 230                 235

GAC TAC TGG GGC CAA GGG ACC ACC GTC ACC GTC TCC TCA TAA              864
Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser  *
240                 245                 250

TAAGCTAGCG GAGCTGCATG CAAATTCTAT TTCAAGGAGA CAGTCATA ATG AAA TAC     921
                                                     Met Lys Tyr
                                                     -22     -20

CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA      969
Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro
                -15                 -10                 -5

GCG ATG GCC GAC ATC GAG CTC ACC CAG TCT CCA TCT TCC ATG TAT GCA     1017
Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Ser Ser Met Tyr Ala
            1               5                   10

TCT CTA GGA GAG AGA ATC ACT ATC ACT TGC AAG GCG AGT CAG GAC ATT     1065
Ser Leu Gly Glu Arg Ile Thr Ile Thr Cys Lys Ala Ser Gln Asp Ile
    15                  20                  25

AAT ACC TAT TTA ACC TGG TTC CAG CAG AAA CCA GGG AAA TCT CCC AAG     1113
Asn Thr Tyr Leu Thr Trp Phe Gln Gln Lys Pro Gly Lys Ser Pro Lys
30                  35                  40                  45

ACC CTG ATC TAT CGT GCA AAC AGA TTG CTA GAT GGG GTC CCA TCA AGG     1161
Thr Leu Ile Tyr Arg Ala Asn Arg Leu Leu Asp Gly Val Pro Ser Arg
                50                  55                  60

TTC AGT GGC AGT GGA TCT GGG CAA GAT TAT TCT CTC ACC ATC AGC AGC     1209
Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu Thr Ile Ser Ser
            65                  70                  75

CTG GAC TAT GAA GAT ATG GGA ATT TAT TAT TGT CTA CAA TAT GAT GAG     1257
Leu Asp Tyr Glu Asp Met Gly Ile Tyr Tyr Cys Leu Gln Tyr Asp Glu
        80                  85                  90

TTG TAC ACG TTC GGA GGG GGG ACC AAG CTC GAG ATC AAA CGG GGT GGA     1305
Leu Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly Gly
    95                  100                 105

GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC GGA GTC GAC ATC GAA     1353
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Asp Ile Glu
110                 115                 120                 125
```

CTC ACT CAG TCT CCA TTC TCC CTG ACT GTG ACA GCA GGA GAG AAG GTC 1401
Leu Thr Gln Ser Pro Phe Ser Leu Thr Val Thr Ala Gly Glu Lys Val
130 135 140

ACT ATG AAT TGC AAG TCC GGT CAG AGT CTG TTA AAC AGT GTA AAT CAG 1449
Thr Met Asn Cys Lys Ser Gly Gln Ser Leu Leu Asn Ser Val Asn Gln
145 150 155

AGG AAC TAC TTG ACC TGG TAC CAG CAG AAG CCA GGG CAG CCT CCT AAA 1497
Arg Asn Tyr Leu Thr Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys
160 165 170

CTG TTG ATC TAC TGG GCA TCC ACT AGG GAA TCT GGA GTC CCT GAT CGC 1545
Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg
175 180 185

TTC ACA GCC AGT GGA TCT GGA ACA GAT TTC ACT CTC ACC ATC AGC AGT 1593
Phe Thr Ala Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
190 195 200 205

GTG CAG GCT GAA GAC CTG GCA GTT TAT TAC TGT CAG AAT GAT TAT ACT 1641
Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn Asp Tyr Thr
210 215 220

TAT CCG TTC ACG TTC GGA GGG GGG ACC AAG CTC GAA ATC AAA CGG TAA 1689
Tyr Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg *
225 230 235

TAAGCGGCCG CGAATTC 1706

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 275 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22 -20 -15 -10

Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp
-5 1 5 10

Leu Val Lys Pro Gly Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly
15 20 25

Phe Thr Phe Ser Ser Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp
30 35 40

Lys Ser Leu Glu Trp Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr
45 50 55

Tyr Tyr Ser Asp Asn Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
60 65 70

Gly Lys Asn Thr Leu Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp
75 80 85 90

Thr Ala Val Tyr Tyr Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr
95 100 105

Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly
110 115 120

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Gly Ser
125 130 135

Ala Gln Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly
140 145 150

Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser
155 160 165 170

Tyr Val Met His Trp Val Lys Gln Lys Pro Gly Gln Gly Leu Glu Trp
175 180 185

Ile Gly Tyr Ile Tyr Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys
190 195 200

Phe Lys Gly Lys Ala Thr Leu Thr Ser Asp Lys Ser Ser Ser Thr Ala
205 210 215

Tyr Met Glu Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr
220 225 230

Cys Ser Arg Arg Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val
235 240 245 250

Ser Ser *

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 259 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
-22 -20 -15 -10

Ala Gln Pro Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Ser Ser
-5 1 5 10

Met Tyr Ala Ser Leu Gly Glu Arg Ile Thr Ile Thr Cys Lys Ala Ser
15 20 25

Gln Asp Ile Asn Thr Tyr Leu Thr Trp Phe Gln Gln Lys Pro Gly Lys
30 35 40

Ser Pro Lys Thr Leu Ile Tyr Arg Ala Asn Arg Leu Leu Asp Gly Val
45 50 55

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu Thr
60 65 70

Ile Ser Ser Leu Asp Tyr Glu Asp Met Gly Ile Tyr Tyr Cys Leu Gln
75 80 85 90

Tyr Asp Glu Leu Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
95 100 105

Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Val
110 115 120

Asp Ile Glu Leu Thr Gln Ser Pro Phe Ser Leu Thr Val Thr Ala Gly
125 130 135

Glu Lys Val Thr Met Asn Cys Lys Ser Gly Gln Ser Leu Leu Asn Ser
140 145 150

Val Asn Gln Arg Asn Tyr Leu Thr Trp Tyr Gln Gln Lys Pro Gly Gln
155 160 165 170

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
175 180 185

Pro Asp Arg Phe Thr Ala Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
190 195 200

Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn
205 210 215

Asp Tyr Thr Tyr Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
220 225 230

Lys Arg *
235

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CACCATCTCC AGAGACAATG GCAAG 25

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GAGCGCGAGC TCGGCCGAAC CGGCCGATCC GCCACCGCCA GAGCC 45

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CAGGATCCGG CCGGTTCGGC CCAGGTCCAG CTGCAACAGT CAGGA 45

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CTACATGAAT TCGCTAGCTT ATTATGAGGA GACGGTGACG GTGGTCCCTT GGC 53

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TAATAAGCTA GCGGAGCTGC ATGCAAATTC TATTTC 36

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ACCAAGCTCG AGATCAAACG GGG 23

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AATGTCGAAT TCGTCGACTC CGCCACCGCC AGAGCC 36

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ATTGGAGTCG ACATCGAACT CACTCAGTCT CCATTCTCC 39

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TGAAGTGAAT TCGCGGCCGC TTATTACCGT TTGATTTCGA GCTTGGTCCC 50

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CGAATTCGGT CACCGTCTCC TCACAGGTCC AGTTGCAACA G 41

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CGAATTCTCG AGATCAAACG GGACATCGAA CTCACTCAGT CTCC 44

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer DBL.12

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CGAATTCGGT CACCGTCTCC TCACAGGTGC AGTTGCAGGA G 41

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer PCR.51

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

AGGTSMAMCT GCAGSAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer PCR.89

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

TGAGGAGACG GTGACCGTGG TCCCTTGGCC CC 32

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer PCR.90

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GACATTGAGC TCACCCAGTC TCCA 24

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:
(B) CLONE: primer PCR.116

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

GTTAGATCTC GAGCTTGGTC CC 22

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

```
Thr Thr Val Thr Val Ser Ser Gln Val Gln Leu Gln Gln
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```
Lys Leu Glu Ile Lys Arg Asp Ile Glu Leu Thr Gln
1               5                   10
```

BP/A/II/12
page 14

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

INTERNATIONAL FORM

To Dr. P van der Logt
Unilever Research
Colworth Laboratory
Biosciences Division
Colworth House, Sharnbrook
Bedford MK44 1LQ

NAME AND ADDRESS OF DEPOSITOR

RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT issued pursuant to Rule 7.1 by the INTERNATIONAL DEPOSITARY AUTHORITY identified at the bottom of this page

I. IDENTIFICATION OF THE MICROORGANISM	
Identification reference given by the DEPOSITOR: Escherichia coli ScFv 4715.myc	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: NCTC 12916

II. SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION

The microorganism identified under I above was accompanied by:

[ ] a scientific description

[X] a proposed taxonomic designation

(Mark with a cross where applicable)

III. RECEIPT AND ACCEPTANCE

This International Depositary Authority accepts the microorganism identified under I above, which was received by it on 14th October 1995 (date of original deposit)[1]

IV. INTERNATIONAL DEPOSITARY AUTHORITY	
Name: National Collection of Type Cultures Address: Central Public Health Laboratory 61 Colindale Avenue London NW9 5HT	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Date: 4/12/95 Barry Holmes B Holmes Clinical Scientist

1 Where Rule 6.4(d) applies, such date is the date on which the status of international depositary authority was acquired; where a deposit made outside the Budapest Treaty after the acquisition of the status of international depositary authority is converted into a deposit under the Budapest Treaty, such date is the date on which the microorganism was received by the international depositary authority.

Form BP/4 (sole page)

BP/4/II/12
page 24

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

INTERNATIONAL FORM

Dr P van der Logt
..ilever Research
..lworth Laboratory
..osciences Division
..lworth House, Sharnbrook
..dford MK44 1LQ

..ME AND ADDRESS OF THE PARTY WHOM THE VIABILITY STATEMENT IS ISSUED

VIABILITY STATEMENT
issued pursuant to Rule 10.2 by the INTERNATIONAL DEPOSITARY AUTHORITY identified on the following page

I. DEPOSITOR	II. IDENTIFICATION OF THE MICROORGANISM
...: Dr P van der Logt ....: As above	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: NCTC 12916 Date of the deposit or of the transfer: 14th October 1995

..VIABILITY STATEMENT

.. viability of the microorganism identified under II above was tested ..[2]. On that date, the said microorganism was

[3] viable

no longer viable

..Indicate the date of the original deposit or, where a new deposit or a transfer has been made, the most recent relevant date (date of the new deposit or date of the transfer).

..In the cases referred to in Rule 10.2(a)(ii) and (iii), refer to the most recent viability test.

..Mark with a cross the applicable box.

..m BP/9 (first page)

BP/A/II/12
page 25

IV. CONDITIONS UNDER WHICH THE VIABILITY TEST HAS BEEN PERFORMED[4]

Nutrient Agar without additions (a)
Agar (17g Bacto-Tryptone, 10g Bacto-Yeast Extract, 5g NaCl, per litre) with 100ug/ml ampicillin and 1% glucose (b)
Aerobic incubation, 37°C, 24 hrs
(a) 8 x 10^5 cfu/ml
(b) 6 x 10^5 cfu/ml

INTERNATIONAL DEPOSITARY AUTHORITY	
Name: National Collection of Type Cultures Central Public Health Laboratory, Address: 61 Colindale Avenue, London NW9 5HT Telephone: 0181-200 4400 Telex: 8953942 ([illegible] G) Fax: 0181-200 7874	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Barry Holmes Date: 4/12/95 B Holmes Clinical Scientist

[4] Fill in if the information has been requested and if the results of the test were negative.

BP/A/II/12
page 14

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

INTERNATIONAL FORM

TO Dr. P van der Logt
Unilever Research
Colworth Laboratory
Biosciences Division
Colworth House, Sharnbrook
Bedford MK44 1LQ

NAME AND ADDRESS OF DEPOSITOR

RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT issued pursuant to Rule 7.1 by the INTERNATIONAL DEPOSITARY AUTHORITY identified at the bottom of this page

I. IDENTIFICATION OF THE MICROORGANISM

Identification reference given by the DEPOSITOR:	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY:
Escherichia coli Fv 3418	NCTC 12915

II. SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION

The microorganism identified under I above was accompanied by:

☐ a scientific description

☒ a proposed taxonomic designation

(Mark with a cross where applicable)

III. RECEIPT AND ACCEPTANCE

This International Depositary Authority accepts the microorganism identified under I above, which was received by it on 14th October 1995 (date of original deposit)[1]

IV. INTERNATIONAL DEPOSITARY AUTHORITY

Name: National Collection of Type Cultures Address: Central Public Health Laboratory 61 Colindale Avenue London NW9 5HT	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Barry Holmes Date: 4/12/95 B Holmes Clinical Scientist

1 Where Rule 6.4(d) applies, such date is the date on which the status of international depositary authority was acquired; where a deposit made outside the Budapest Treaty after the acquisition of the status of international depositary authority is converted into a deposit under the Budapest Treaty, such date is the date on which the microorganism was received by the international depositary authority.

Form BP/4 (sole page)

BP/4/II/12
page 24

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

INTERNATIONAL FORM

Dr P van der Logt
Unilever Research
Colworth Laboratory
Biosciences Division
Colworth House, Sharnbrook
Bedford MK44 1LQ

NAME AND ADDRESS OF THE PARTY TO WHOM THE VIABILITY STATEMENT IS ISSUED

VIABILITY STATEMENT
issued pursuant to Rule 10.2 by the
INTERNATIONAL DEPOSITARY AUTHORITY
identified on the following page

I. DEPOSITOR	II. IDENTIFICATION OF THE MICROORGANISM
Name: Dr P van der Logt Address: As above	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: NCTC 12915 Date of the deposit or of the transfer:[1] 14th October 1995

III. VIABILITY STATEMENT

The viability of the microorganism identified under II above was tested on 20th November 1995[2]. On that date, the said microorganism was

[3] X viable

[3] no longer viable

[1] Indicate the date of the original deposit or, where a new deposit or a transfer has been made, the most recent relevant date (date of the new deposit or date of the transfer).

[2] In the cases referred to in Rule 10.2(a)(ii) and (iii), refer to the most recent viability test.

[3] Mark with a cross the applicable box.

Form BP/9 (first page)

page 2

IV. CONDITIONS UNDER WHICH THE VIABILITY TEST HAS BEEN PERFORMED[4]

Nutrient Agar without additions (a)
Agar (17g Bacto-Tryptone, 10g Bacto-Yeast Extract, 5g NaCl, per litre) with 100ug/ml ampicillin and 1% glucose (b)
Aerobic incubation, 37°C, 24 hrs.

(a) 2×10^6 cfu/ml

(b) 1×10^6 cfu/ml

INTERNATIONAL DEPOSITARY AUTHORITY		
Name:	National Collection of Type Cultures	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Barry Holmes
Address:	Central Public Health Laboratory 61 Colindale Avenue, London NW9 5HT Telephone: 0181-200 4400 Telex: 8953942 (DEFEND G) Fax: 0181-200 7874	Date: 4/12/95 B Holmes Clinical Scientist

[4] Fill in if the information has been requested and if the results of the test were negative.

Form BP/9 (second and last page)

C L A I M S

1. **A bispecific** or **bivalent** antibody fragment analogue, which comprises a binding complex containing two polypeptide chains, whereby one polypeptide chain comprises two times a variable domain of a heavy chain (V_H) **in series** and the other polypeptide chain comprises two times a variable domain of a light chain (V_L) **in series**, and the binding complex contains two pairs of variable domains (V_H-A//V_L-A and V_H-B//V_L-B).

2. An antibody fragment analogue according to claim 1, in which one polypeptide chain comprises a first V_H <u>connected to</u> a second V_H and the other polypeptide chain comprises a first V_L <u>connected to</u> a second V_L.

3. An antibody fragment analogue according to claim 2, in which the two V_H's are directly connected to each other without an intermediate peptide linker.

4. An antibody fragment analogue according to claim 2, in which the two V_L's are directly connected to each other without an intermediate peptide linker.

5. An antibody fragment analogue according to claim 3 or claim 4, in which one polypeptide chain comprises a first V_H directly connected to a second V_H, and the other polypeptide chain comprises a first V_L directly connected to a second V_L.

6. An antibody fragment analogue according to claim 2, in which the two V_H's are connected to each other by a peptide <u>linker</u> and also the two V_L's are connected to each other by a peptide linker, each peptide linker comprising at least one amino acid residue.

7. An antibody fragment analogue according to claim 6, in which one polypeptide chain comprises a first V_H (V_H-A) followed by a second V_H (V_H-B) and the other polypeptide chain comprises a first V_L (V_L-A) followed by a second V_L (V_L-B), and in which the two V_H's are connected to each other by a peptide linker (Li_H), thus [V_H-A * Li_H * V_H-B], and also the two V_L's are connected to each other by a peptide linker (Li_L), thus [V_L-A * Li_L * V_L-B], each peptide linker comprising at least 10 amino acid residues.

8. An antibody fragment analogue according to claim 2, in which one polypeptide chain comprises a first V_H (V_H-A) followed by a second V_H (V_H-B) with or without a connecting peptide linker (Li_H), thus [V_H-A * (Li_H) * V_H-B], and the other polypeptide chain comprises a first V_L (V_L-A) preceded by a second V_L (V_L-B) with or without a connecting peptide linker (Li_L), thus [V_L-B * (Li_L) * V_L-A].

9. An antibody fragment analogue according to claim 1, in which the two variable domains are different resulting in a **bispecific** antibody fragment analogue.

10. An antibody fragment analogue according to claim 1, in which the specificities A and B are the same resulting in a **bivalent** antibody fragment analogue.

11. Use of an antibody fragment analogue according to claim 1, in immunoassays including diagnostic techniques, in agglutination assays, in a purification method, for compositions suitable for therapy, or in other methods in which immunoglobulins or fragments thereof can be used.

12. A process for producing an antibody fragment analogue according to any one of claims 1-10, which comprises

(1) transforming a host by incorporating into that host a DNA encoding the two V_H's in series with or without a

connecting peptide linker and a DNA encoding the two V_L's in series with or without a connecting peptide linker,

(2) culturing such transformed host under conditions whereby the connected V_H's and the connected V_L's are formed, and

(3) allowing the two connected V_H's and the two connected V_L's to combine to each other under formation of a double head antibody fragment analogue, and

(4) optionally collecting the double head antibody fragment analogue.

13. A process for producing an antibody fragment analogue according to any one of claims 1-10, which comprises

(1) transforming a first host by incorporating into that first host a DNA encoding the two V_H's in series with or without a connecting peptide linker,

(2) transforming a second host by incorporating into that second host a DNA encoding the two V_L's in series with or without a connecting peptide linker,

(3) culturing the first and second transformed host under conditions whereby the connected V_H's and the connected V_L's, respectively, are formed,

(4) optionally collecting separately the two connected V_H's and the two connected V_L's, and

(5) combining the connected V_H's and the connected V_L's under conditions that they can form a double head antibody fragment analogue.

14. A process according to claim 12 or 13, in which the host is selected from the group consisting of prokaryotic micro-organisms comprising Gram-negative bacteria (e.g. *E. coli*) and Gram-positive bacteria (e.g. *B. subtilis* or lactic acid bacteria), lower eukaryotic microorganisms comprising yeasts (e.g. belonging to the genera *Saccharomyces*, *Kluyveromyces*, *Pichia*, and *Hansenula*) and moulds (e.g. belonging to the genera *Aspergillus*,

Neurospora or *Trichoderma*), and higher eukaryotic organisms (e.g. plants) or cell cultures thereof (e.g. hybridoma's).

Fig.1.

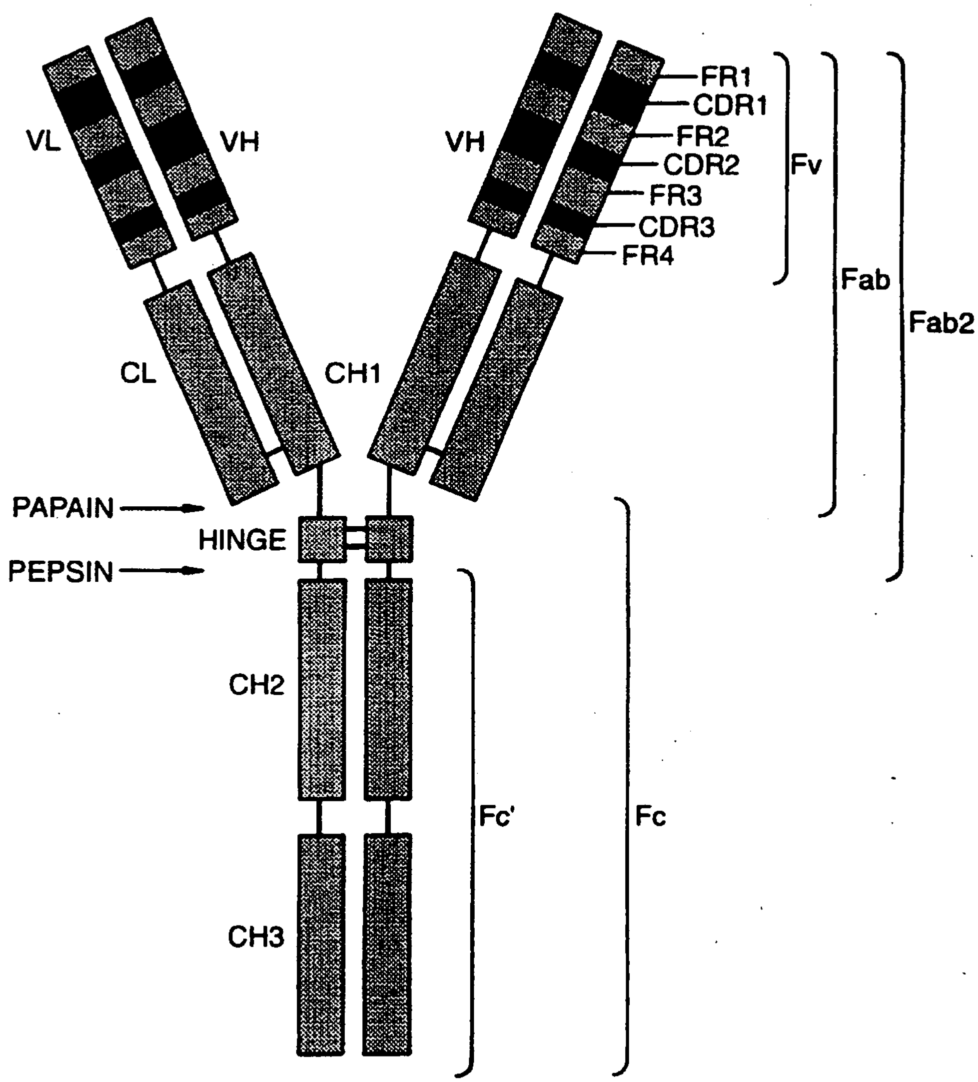

VL
VH
VH
FR1
CDR1
FR2
CDR2
FR3
CDR3
FR4
Fv
Fab
Fab2
CL
CH1
PAPAIN
HINGE
PEPSIN
CH2
CH3
Fc'
Fc

Fig.2A.

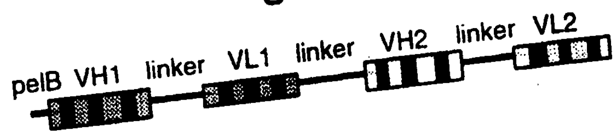


Fig.2B.

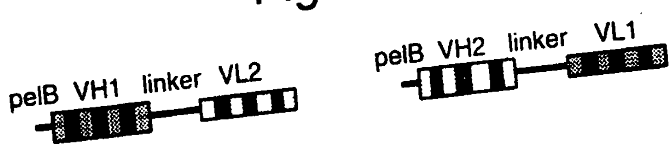
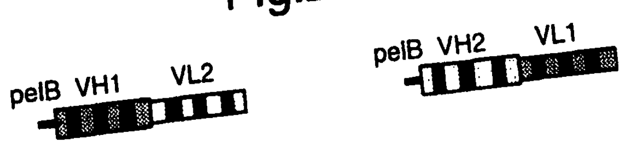


Fig.2C.



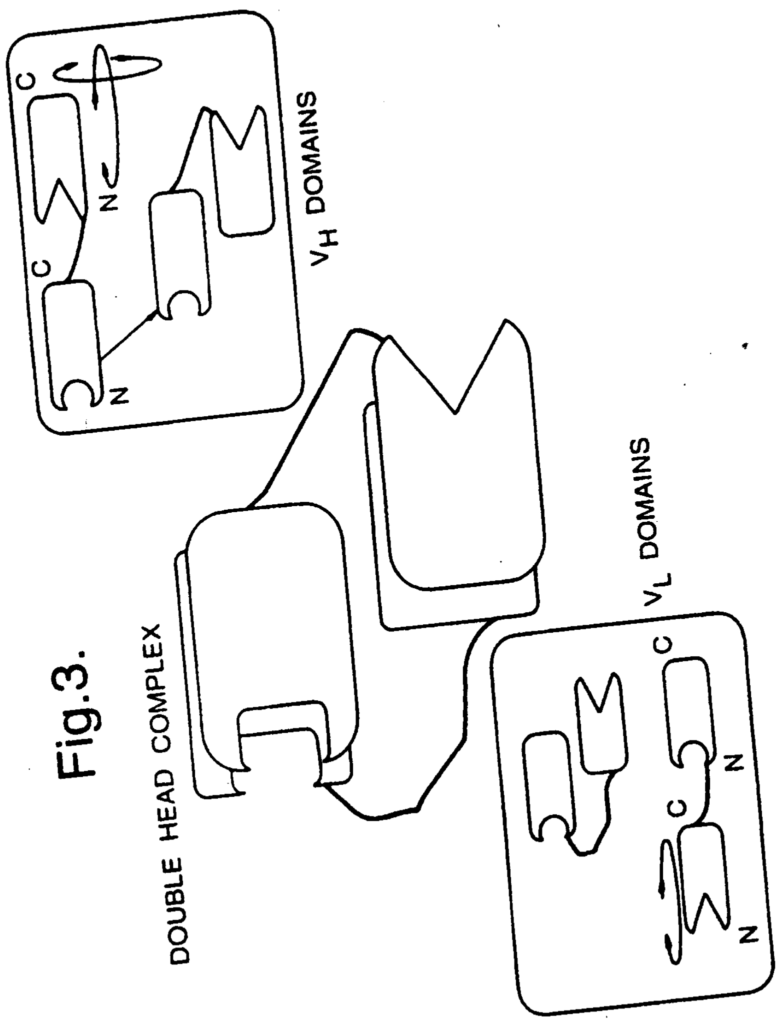


Fig.3.

Fig.4.

```
                 D  I  E  L  T  Q  S  P  A  S  L  S  A  S  V  G  E
GAATTCGGCCGACATCGAGCTCACCCAGTCTCCAGCCTCCCTTTCTGCGTCTGTGGGAGA  60
EcoRI           SacI

 T  V  T  I  T  C  R  A  S  G  N  I  H  N  Y  L  A  W  Y  Q
AACTGTCACCATCACATGTCGAGCAAGTGGGAATATTCACAATTATTTAGCATGGTATCA 120

 Q  K  Q  G  K  S  P  Q  L  L  V  Y  Y  T  T  T  L  A  D  G
GCAGAAACAGGGAAAATCTCCTCAGCTCCTGGTCTATTATACAACAACCTTAGCAGATGG 180

 V  P  S  R  F  S  G  S  G  S  G  T  Q  Y  S  L  K  I  N  S
TGTGCCATCAAGGTTCAGTGGCAGTGGATCAGGAACACAATATTCTCTCAAGATCAACAG 240

 L  Q  P  E  D  F  G  S  Y  Y  C  Q  H  F  W  S  T  P  R  T
CCTGCAACCTGAAGATTTTGGGAGTTATTACTGTCAACATTTTTGGAGTACTCCTCGGAC 300

 F  G  G  G  T  K  L  E  I  K  R  G  G  G  G  S  G  G  G  G
GTTCGGTGGAGGGACCAAGCTCGAGATCAAACGGGGTGGAGGCGGTTCAGGCGGAGGTGG 360
                    XhoI

 S  G  G  G  G  S  Q  V  Q  L  Q  E  S  G  P  G  L  V  A  P
CTCTGGCGGTGGCGGATCGCAGGTGCAGCTGCAGGAGTCAGGACCTGGCCTGGTGGCGCC 420
                            PstI

 S  Q  S  L  S  I  T  C  T  V  S  G  F  S  L  T  G  Y  G  V
CTCACAGAGCCTGTCCATCACATGCACCGTCTCAGGGTTCTCATTAACCGGCTATGGTGT 480

 N  W  V  R  Q  P  P  G  K  G  L  E  W  L  G  M  I  W  G  D
AAACTGGGTTCGCCAGCCTCCAGGAAAGGGTCTGGAGTGGCTGGGAATGATTTGGGGTGA 540

 G  N  T  D  Y  N  S  A  L  K  S  R  L  S  I  S  K  D  N  S
TGGAAACACAGACTATAATTCAGCTCTCAAATCCAGACTGAGCATCAGCAAGGACAACTC 600

 K  S  Q  V  F  L  K  M  N  S  L  H  T  D  D  T  A  R  Y  Y
CAAGAGCCAAGTTTTCTTAAAAATGAACAGTCTGCACACTGATGACACAGCCAGGTACTA 660

 C  A  R  E  R  D  Y  R  L  D  Y  W  G  Q  G  T  T  V  T  V
CTGTGCCAGAGAGAGAGATTATAGGCTTGACTACTGGGGCCAAGGGACCACGGTCACCGT 720

 S  S  *
CTCCTCATGATAAGCTT
            HindIII 737
```

DIELTQSPAS---GGGTKLEIKR = VL*Lys*
GGGGSGGGGSGGGGS = Linker
QVQLQESGPG---GQGTTVTVSS = VH*Lys*

Fig.5.

```
                                        M  K  Y  L  L  P  T  A
pelB      AAGCTTGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAG
LEADER     A  A  G  L  L  L  L  A  A  Q  P  A  M  A  Q  V  Q  L  Q  Q
          CCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAGCAGT
           S  G  P  E  L  V  K  P  G  A  S  V  K  M  S  C  K  A  S  G
          CAGGACCTGAGCTGGTAAAGCCTGGGGCTTCAGTGAAGATGTCCTGCAAGGCTTCTGGAT
           Y  T  F  T  S  Y  V  M  H  W  V  K  Q  K  P  G  Q  G  L  E
          ACACATTCACTAGCTATGTTATGCACTGGGTGAAACAGAAGCCTGGGCAGGGCCTTGAGT
           W  I  G  Y  I  Y  P  Y  N  D  G  T  K  Y  N  E  K  F  K  G
VH3418    GGATTGGATATATTTATCCTTACAATGATGGTACTAAGTACAATGAGAAGTTCAAAGGCA
           K  A  T  L  T  S  D  K  S  S  S  T  A  Y  M  E  L  S  S  L
          AGGCCACACTGACTTCAGACAAATCCTCCAGCACAGCCTACATGGAGCTCAGCAGCCTGA
           T  S  E  D  S  A  V  Y  Y  C  S  R  R  F  D  Y  W  G  Q  G
          CCTCTGAGGACTCTGCGGTCTATTACTGTTCAAGACGCTTTGACTACTGGGGCCAAGGGA
           T  T  V  T  V  S  S
          CCACGGTCACCGTCTCCTCATAATAAGAGCTATGGGAGCTTGCATGCAAATTCTATTTCA
                         M  K  Y  L  L  P  T  A  A  A  G  L  L  L  L
          AGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCG
pelB       A  A  Q  P  A  M  A  D  I  E  L  T  Q  S  P  S  S  M  Y  A
LEADER    CTGCCCAACCAGCGATGGCCGACATCGAGCTCACCCAGTCTCCATCTTCCATGTATGCAT
           S  L  G  E  R  I  T  I  T  C  K  A  S  Q  D  I  N  T  Y  L
          CTCTAGGAGAGAGAATCACTATCACTTGCAAGGCGAGTCAGGACATTAATACCTATTTAA
           T  W  F  Q  Q  K  P  G  K  S  P  K  T  L  I  Y  R  A  N  R
          CCTGGTTCCAGCAGAAACCAGGGAAATCTCCCAAGACCCTGATCTATCGTGCAAACAGAT
           L  L  D  G  V  P  S  R  F  S  G  S  G  S  G  Q  D  Y  S  L
VL3418    TGCTAGATGGGGTCCCATCAAGGTTCAGTGGCAGTGGATCTGGGCAAGATTATTCTCTCA
           T  I  S  S  L  D  Y  E  D  M  G  I  Y  Y  C  L  Q  Y  D  E
          CCATCAGCAGCCTGGACTATGAAGATATGGGAATTTATTATTGTCTACAATATGATGAGT
           L  Y  T  F  G  G  G  T  K  L  E  I  K  R
          TGTACACGTTCGGAGGGGGGACCAAGCTCGAGATCAAACGGTAATAATGATCAAACGGT
          ATAAGGATCCAGCTCGAATTC
```

Fig.6.

pelB LEADER

M K Y L L P T
AAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

A A A G L L L L A A Q P A M A Q V Q L Q
GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G G D L V K P G G S L T L S C A T S
GAGTCAGGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT

G F T F S S Y A F S W V R Q T S D K S L
GGATTCACTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG

VH4715

E W V A T I S S T D T Y T Y Y S D N V K
GAGTGGGTCGCAACCATCAGTAGTACTGATACTTATACCTATTATTCAGACAATGTGAAG

Q R F T I S R D N G K N T L Y L Q M S S
GGGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAATGAGCAGT

L K S E D T A V Y Y C A R H G Y Y G K G
CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC

Y F D Y W G Q G T T V T V S S
TATTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCATAATAAGAGCTATGG

pelB LEADER

M K Y L L P T
GAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

A A A G L L L L A A Q P A M A D I E L T
GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCGACATCGAGCTCACT

Q S P F S L T V T A G E K V T M N C K S
CAGTCTCCATTCTCCCTGACTGTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCC

G Q S L L N S V N Q R N Y L T W Y Q Q K
GGTCAGAGTCTGTTAAACAGTGTAAATCAGAGGAACTACTTGACCTGGTACCAGCAGAAG

P G Q P P K L L I Y W A S T R E S G V P
CCAGGGCAGCCTCCTAAACTGTTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCT

VL4715

D R F T A S G S G T D F T L T I S S V Q
GATCGCTTCACAGCCAGTGGATCTGGAACAGATTTCACTCTCACCATCAGCAGTGTGCAG

A E D L A V Y Y C Q N D Y T Y P F T F G
GCTGAAGACCTGGCAGTTTATTACTGTCAGAATGATTATACTTATCCGTTCACGTTCGGA

Myc-tag

G G T K L E I K R E Q K L I S E E D L N
GGGGGGACCAAGCTCGAGATCAAACGGGAACAAAAACTCATCTCAGAAGAGGATCTGAAT

TAATAAGATCAAACGGTAATAAGGATCCAGCTCGAATTC

Fig.7.

pelB LEADER

M K Y L L P T
AAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

A A A G L L L L A A Q P A M A Q V Q L Q
GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G G D L V K P G G S L T L S C A T S
GAGTCAGGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT

G F T F S S Y A F S W V R Q T S D K S L
GGATTCACTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG

VH4715

E W V A T I S S T D T Y T Y Y S D N V K
GAGTGGGTCGCAACCATCAGTAGTACTGATACTTATACCTATTATTCAGACAATGTGAAG

G R F T I S R D N G K N T L Y L Q M S S
GGGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAATGAGCAGT

L K S E D T A V Y Y C A R H G Y Y G K G
CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC

Y F D Y W G Q G T T V T V S S G G G G S
TATTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGAGGCGGTTCA

LINKER

G G G G S G G G G S D I E L T Q S P F S
GGCGGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCACTCAGTCTCCATTCTCC

L T V T A G E K V T M N C K S G Q S L L
CTGACTGTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCCGGTCAGAGTCTGTTA

N S V N Q R N Y L T W Y Q Q K P G Q P P
AACAGTGTAAATCAGAGGAACTACTTGACCTGGTACCAGCAGAAGCCAGGGCAGCCTCCT

VL4715

K L L I Y W A S T R E S G V P D R F T A
AAACTGTTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCTGATCGCTTCACAGCC

S G S G T D F T L T I S S V Q A E D L A
AGTGGATCTGGAACAGATTTCACTCTCACCATCAGCAGTGTGCAGGCTGAAGACCTGGCA

V Y Y C Q N D Y T Y P F T F G G G T K L
GTTTATTACTGTCAGAATGATTATACTTATCCGTTCACGTTCGGAGGGGGGACCAAGCTC

Myc-tag

E I K R E Q K L I S E E D L N
GAGATCAAACGGGAACAAAAACTCATCTCAGAAGAGGATCTGAATTAATAAGATCAAACG

GTAATAAGGATCCAGCTCGAATTC

8/45

Fig.8.1(2)

M K Y L L P T
AAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

pelB LEADER
A A A G L L L L A A Q P A M A Q V Q L Q
GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G G D L V K P G G S L T L S C A T S
GAGTCAGGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT

VH4715
G F T F S S Y A F S W V R Q T S D K S L
GGATTCACTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG

E W V A T I S S T D T Y T Y Y S D N V K
GAGTGGGTCGCAACCATCAGTAGTACTGATACTTATACCTATTATTCAGACAATGTGAAG

G R F T I S R D N G K N T L Y L Q M S S
GGGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAATGAGCAGT

L K S E D T A V Y Y C A R H G Y Y G K G
CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC

Y F D Y W G Q G T T V T V S S G G G G S
TATTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGAGGCGGTTCA

LINKER
G G G G S G G G G S A G S A Q V Q L Q Q
GGCGGAGGTGGCTCTGGCGGTGGCGGATCGGCCGGTTCGGCCCAGGTCCAGCTGCAACAG

S G P E L V K P G A S V K M S C K A S G
TCAGGACCTGAGCTGGTAAAGCCTGGGGCTTCAGTGAAGATGTCCTGCAAGGCTTCTGGA

Y T F T S Y V M H W V K Q K P G Q G L E
TACACATTCACTAGCTATGTTATGCACTGGGTGAAACAGAAGCCTGGGCAGGGCCTTGAG

VH3418
W I G Y I Y P Y N D G T K Y N E K F K G
TGGATTGGATATATTTATCCTTACAATGATGGTACTAAGTACAATGAGAAGTTCAAAGGC

K A T L T S D K S S S T A Y M E L S S L
AAGGCCACACTGACTTCAGACAAATCCTCCAGCACAGCCTACATGGAGCTCAGCAGCCTG

T S E D S A V Y Y C S R R F D Y W G Q G
ACCTCTGAGGACTCTGCGGTCTATTACTGTTCAAGACGCTTTGACTACTGGGGCCAAGGG

T T V T V S S
ACCACCGTCACCGTCTCCTCATAATAAGCTAGCGGAGCTGCATGCAAATTCTATTTCAAG

pelB LEADER
M K Y L L P T A A A G L L L L A
GAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCT

A Q P A M A D I E L T Q S P S S M Y A S
GCCCAACCAGCGATGGCCGACATCGAGCTCACCCAGTCTCCATCTTCCATGTATGCATCT

VL3418
L G E R I T I T C K A S Q D I N T Y L T
CTAGGAGAGAGAATCACTATCACTTGCAAGGCGAGTCAGGACATTAATACCTATTTAACC

W F Q Q K P G K S P K T L I Y R A N R L
TGGTTCCAGCAGAAACCAGGGAAATCTCCCAAGACCCTGATCTATCGTGCAAACAGATTG

Fig.8. 2(2)

```
 L  D  G  V  P  S  R  F  S  G  S  G  S  G  Q  D  Y  S  L  T
CTAGATGGGGTCCCATCAAGGTTCAGTGGCAGTGGATCTGGGCAAGATTATTCTCTCACC

 I  S  S  L  D  Y  E  D  M  G  I  Y  Y  C  L  Q  Y  D  E  L
ATCAGCAGCCTGGACTATGAAGATATGGGAATTTATTATTGTCTACAATATGATGAGTTG

 Y  T  F  G  G  G  T  K  L  E  I  K  R| G  G  G  G  S  G  G
TACACGTTCGGAGGGGGGACCAAGCTCGAGATCAAACGGGGTGGAGGCGGTTCAGGCGGA
LINKER
 G  G  S  G  G  G  G  V| D  I  E  L  T  Q  S  P  F  S  L  T
GGTGGCTCTGGCGGTGGCGGAGTCGACATCGAACTCACTCAGTCTCCATTCTCCCTGACT

 V  T  A  G  E  K  V  T  M  N  C  K  S  G  Q  S  L  L  N  S
GTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCCGGTCAGAGTCTGTTAAACAGT

 V  N  Q  R  N  Y  L  T  W  Y  Q  Q  K  P  G  Q  P  P  K  L
GTAAATCAGAGGAACTACTTGACCTGGTACCAGCAGAAGCCAGGGCAGCCTCCTAAACTG

VL4715
 L  I  Y  W  A  S  T  R  E  S  G  V  P  D  R  F  T  A  S  G
TTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCTGATCGCTTCACAGCCAGTGGA

 S  G  T  D  F  T  L  T  I  S  S  V  Q  A  E  D  L  A  V  Y
TCTGGAACAGATTTCACTCTCACCATCAGCAGTGTGCAGGCTGAAGACCTGGCAGTTTAT

 Y  C  Q  N  D  Y  T  Y  P  F  T  F  G  G  G  T  K  L  E  I
TACTGTCAGAATGATTATACTTATCCGTTCACGTTCGGAGGGGGGACCAAGCTCGAAATC

 K  R
AAACGGTAATAAGCGGCCGCGAATTC
```

Fig.9.

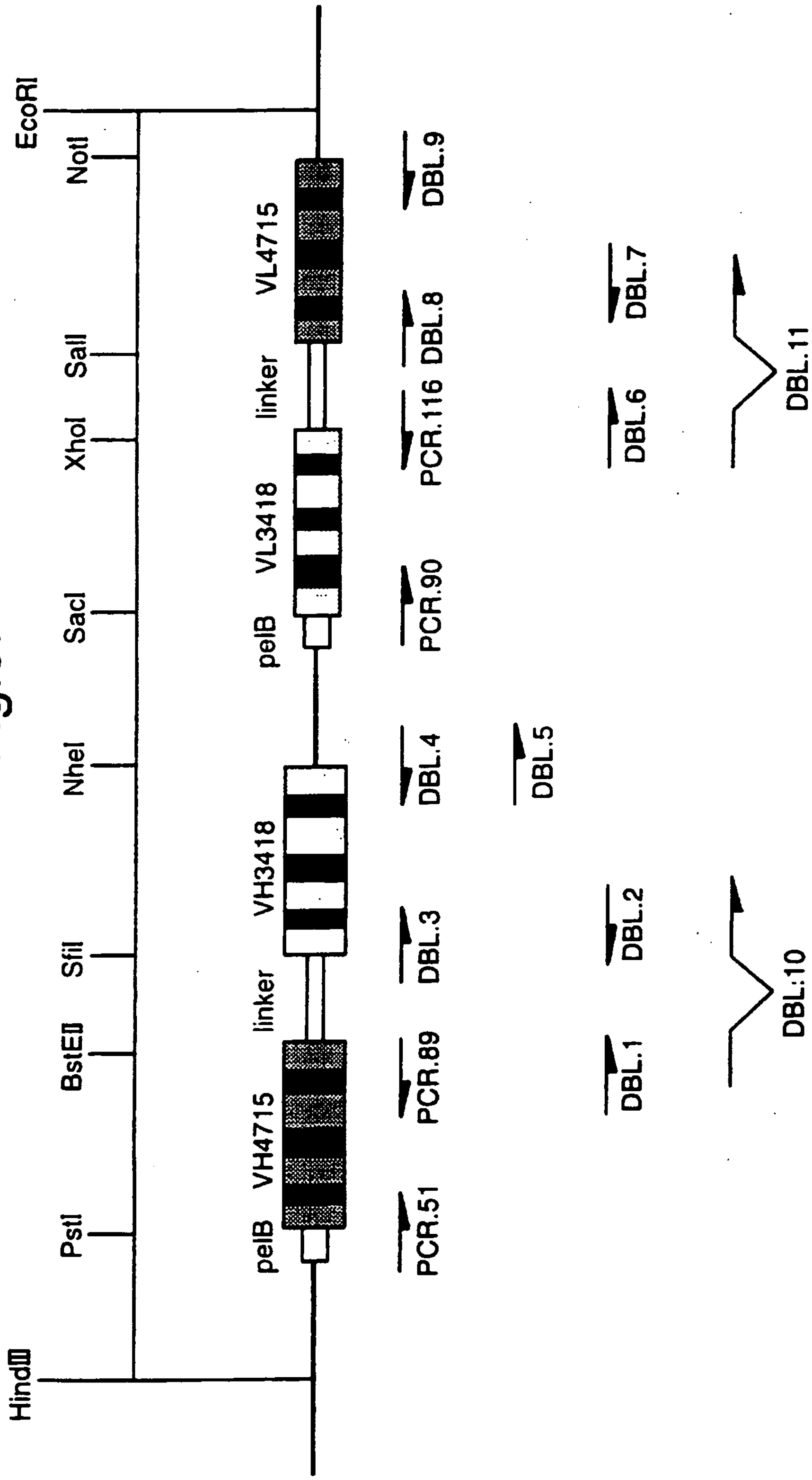

EcoRI
NotI
SalI
XhoI
SacI
NheI
SfiI
BstEII
PstI
HindIII
VL4715
linker
VL3418
pelB
VH3418
linker
VH4715
pelB
DBL.9
DBL.8
PCR.116
PCR.90
DBL.4
DBL.3
PCR.89
PCR.51
DBL.5
DBL.7
DBL.6
DBL.2
DBL.1
DBL.11
DBL:10

Fig.10.

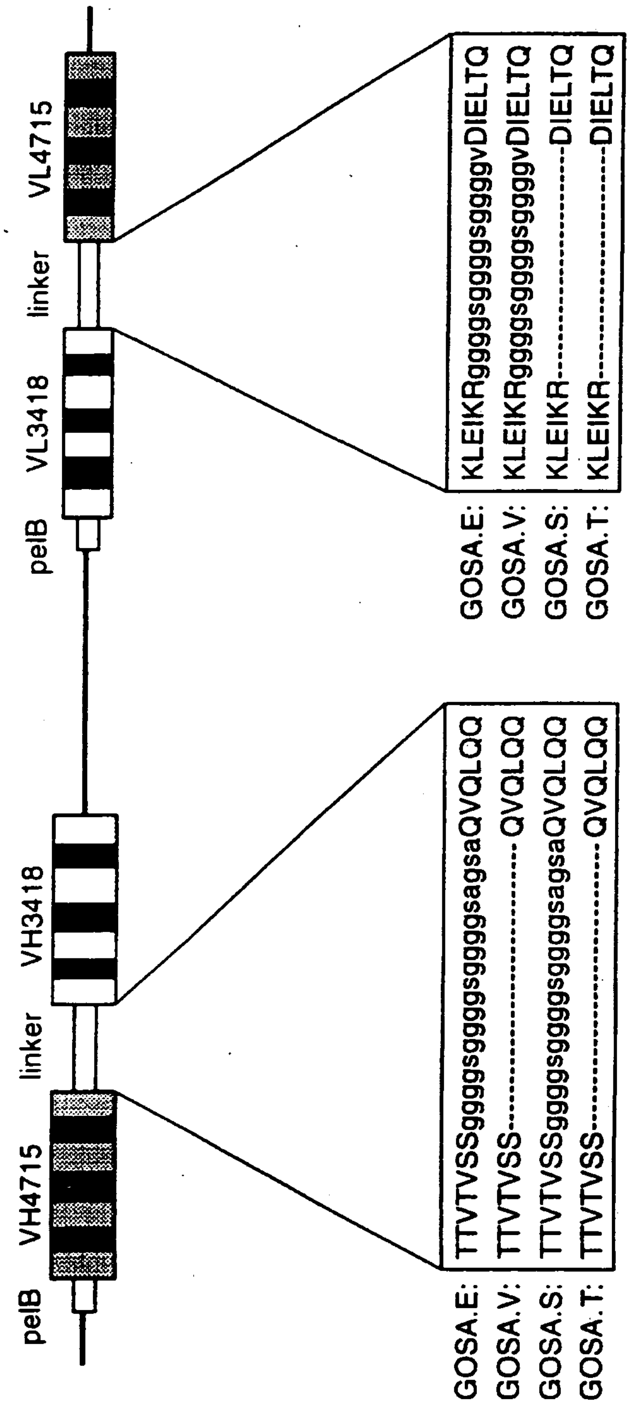

pelB
VH4715
linker
VH3418
pelB
VL3418
linker
VL4715
GOSA.E: TTVTVSSggggsggggsggggsagsaQVQLQQ
GOSA.V: TTVTVSS------------------QVQLQQ
GOSA.S: TTVTVSSggggsggggsggggsagsaQVQLQQ
GOSA.T: TTVTVSS------------------QVQLQQ
GOSA.E: KLEIKRggggsggggsggggvDIELTQ
GOSA.V: KLEIKRggggsggggsggggvDIELTQ
GOSA.S: KLEIKR---------------DIELTQ
GOSA.T: KLEIKR---------------DIELTQ

Fig.11.

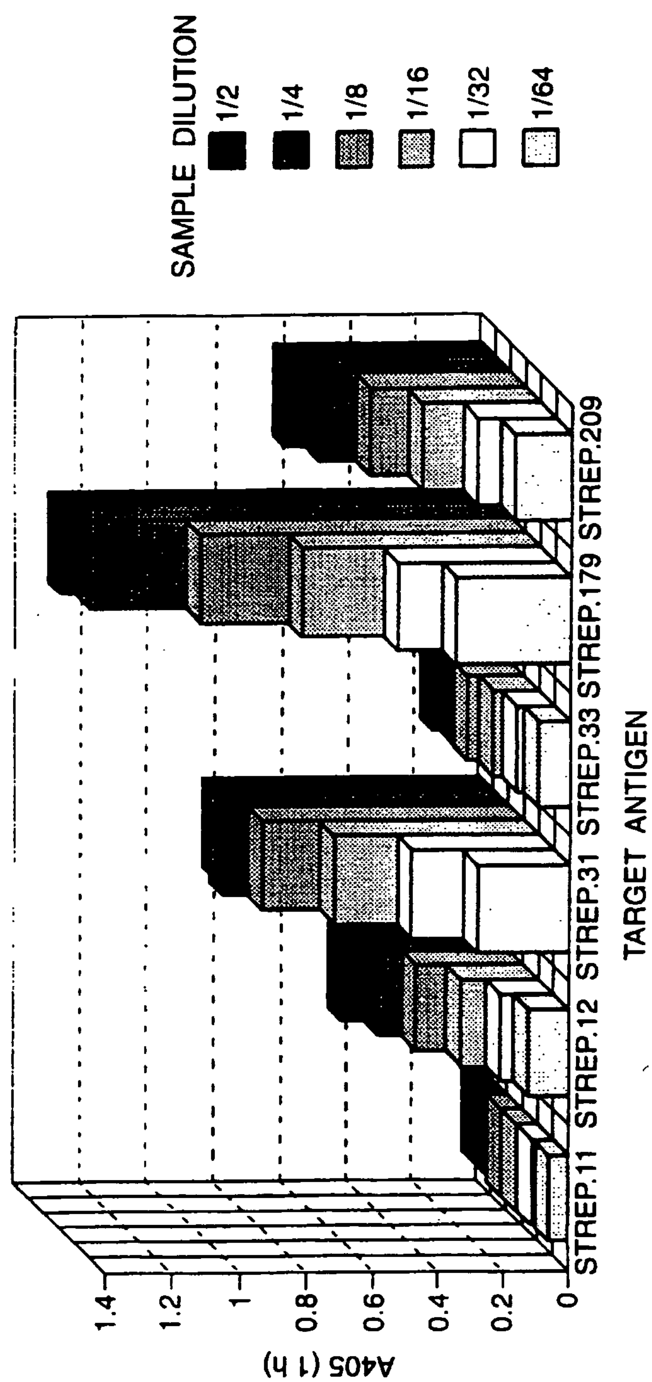

SAMPLE DILUTION
1/2
1/4
1/8
1/16
1/32
1/64
A405 (1 h)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
STREP.11
STREP.12
STREP.31
STREP.33
STREP.179
STREP.209
TARGET ANTIGEN

Fig.12.

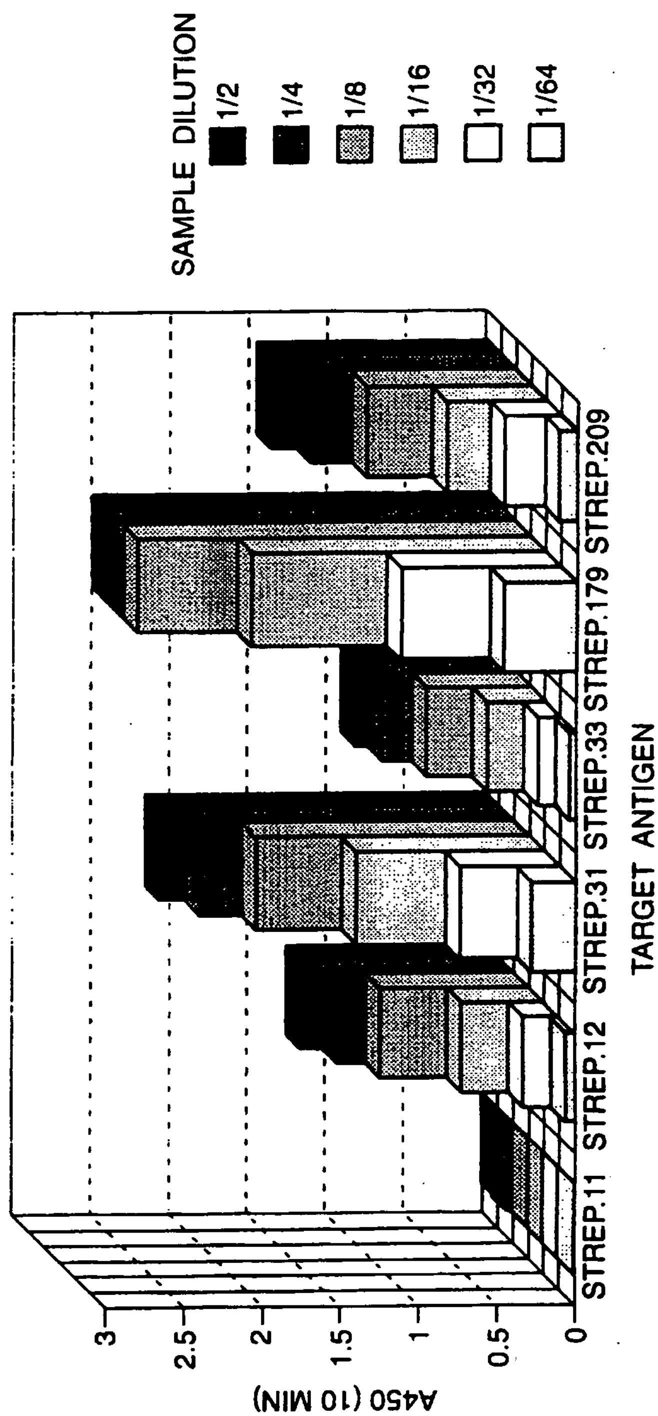

SAMPLE DILUTION
1/2
1/4
1/8
1/16
1/32
1/64
A450 (10 MIN)
0
0.5
1
1.5
2
2.5
3
STREP.11
STREP.12
STREP.31
STREP.33
STREP.179
STREP.209
TARGET ANTIGEN

Fig.13.

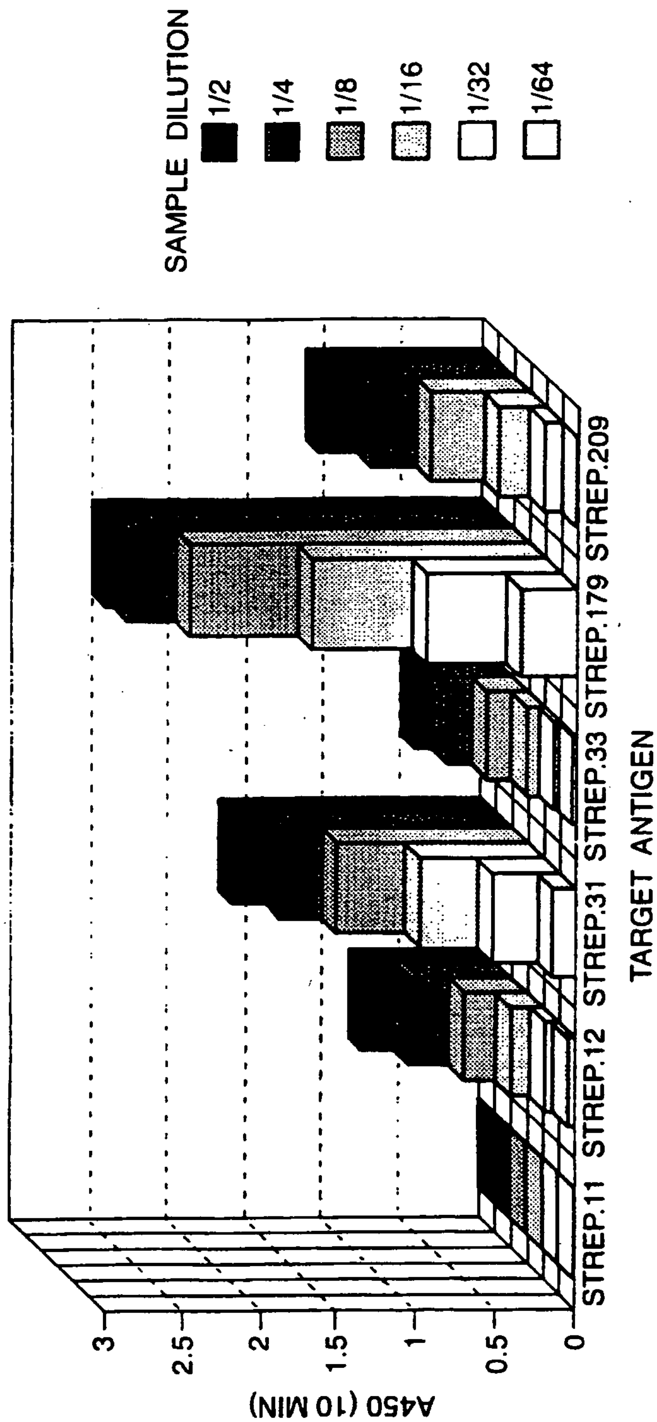

SAMPLE DILUTION
1/2
1/4
1/8
1/16
1/32
1/64
A450 (10 MIN)
3
2.5
2
1.5
1
0.5
0
STREP.11
STREP.12
STREP.31
STREP.33
STREP.179
STREP.209
TARGET ANTIGEN

Fig.14.

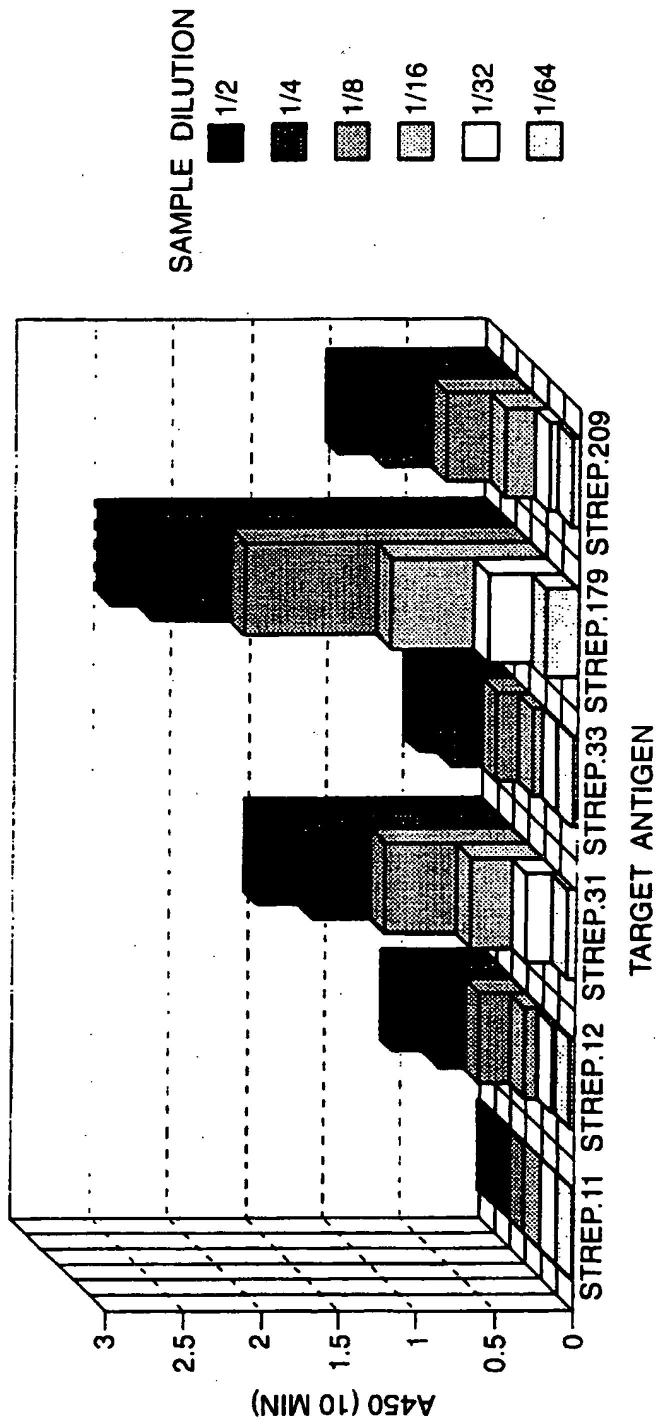

SAMPLE DILUTION
1/2
1/4
1/8
1/16
1/32
1/64
A450 (10 MIN)
0
0.5
1
1.5
2
2.5
3
STREP.11
STREP.12
STREP.31
STREP.33
STREP.179
STREP.209
TARGET ANTIGEN

Fig.15.

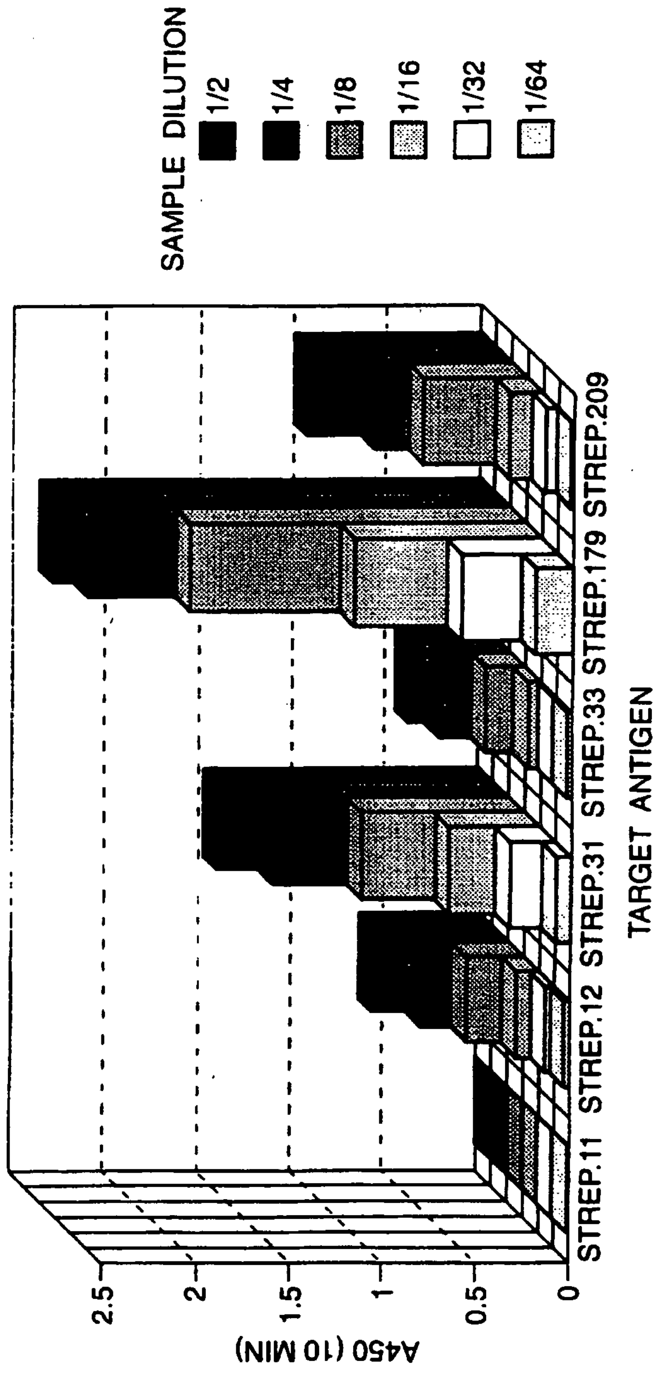

SAMPLE DILUTION
1/2
1/4
1/8
1/16
1/32
1/64
A450 (10 MIN)
0
0.5
1
1.5
2
2.5
STREP.11
STREP.12
STREP.31
STREP.33
STREP.179
STREP.209
TARGET ANTIGEN

Fig.16.

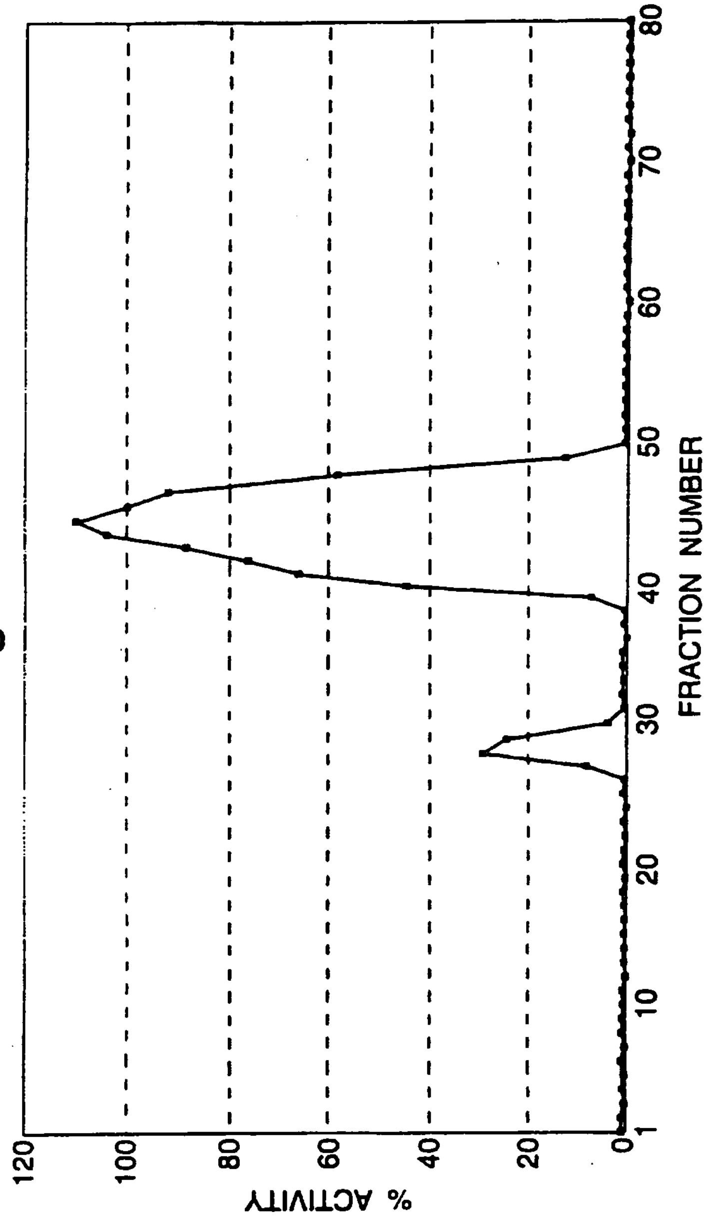
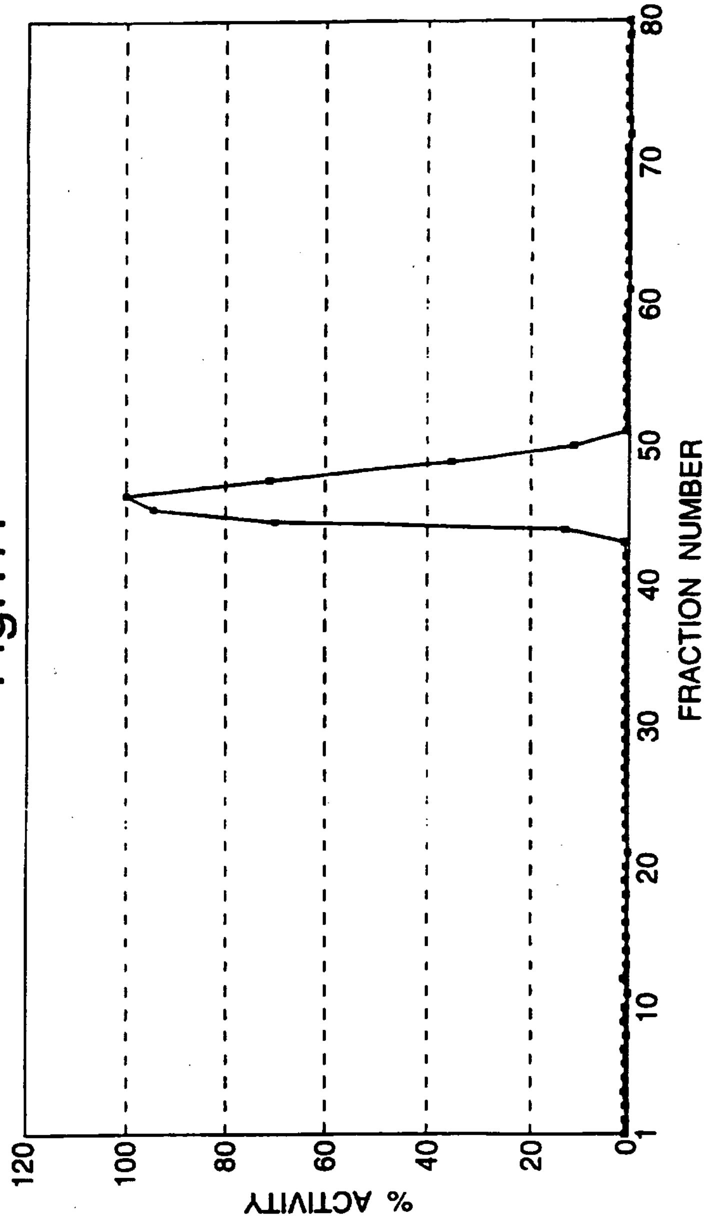

% ACTIVITY
120
100
80
60
40
20
0
FRACTION NUMBER
1
10
20
30
40
50
60
70
80

Fig.17.


% ACTIVITY
120
100
80
60
40
20
0
FRACTION NUMBER
1
10
20
30
40
50
60
70
80

19/45

NOT TO BE TAKEN
INTO CONSIDERATION
FOR THE PURPOSES
OF INTERNATIONAL PROCESSING

(See Section 310(d)(ii) of the Administrative Instructions)

Fig.19.

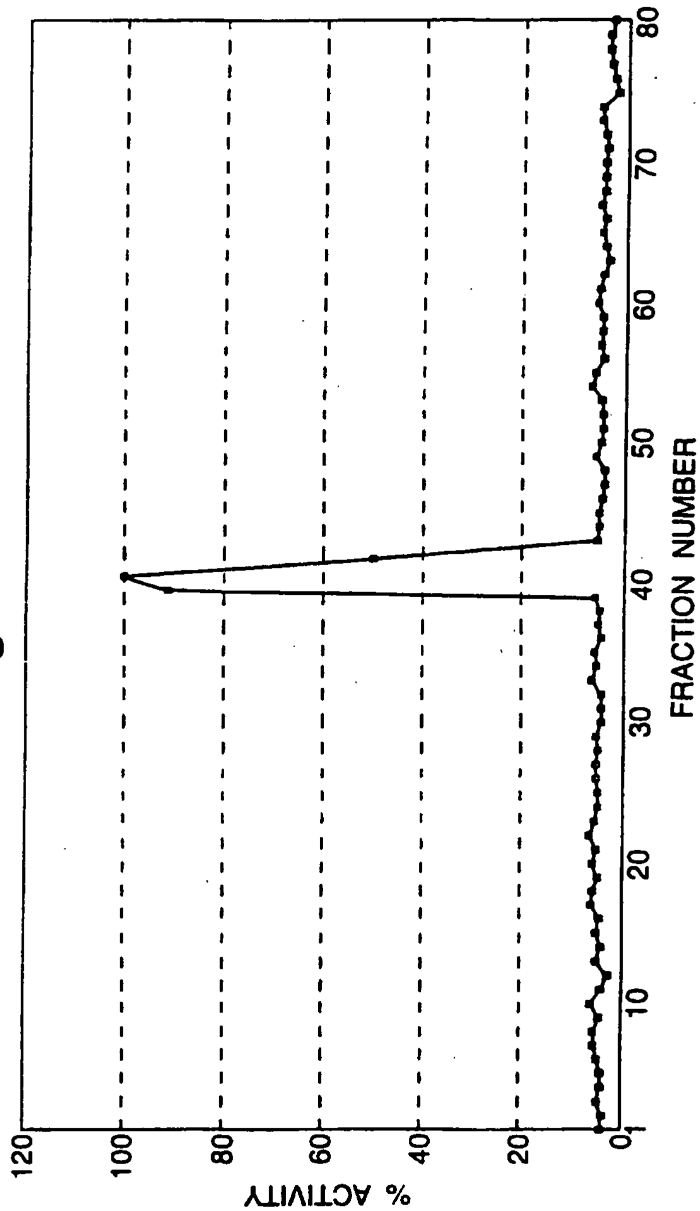

% ACTIVITY
120
100
80
60
40
20
0
0
10
20
30
40
50
60
70
80
FRACTION NUMBER

Fig.21.

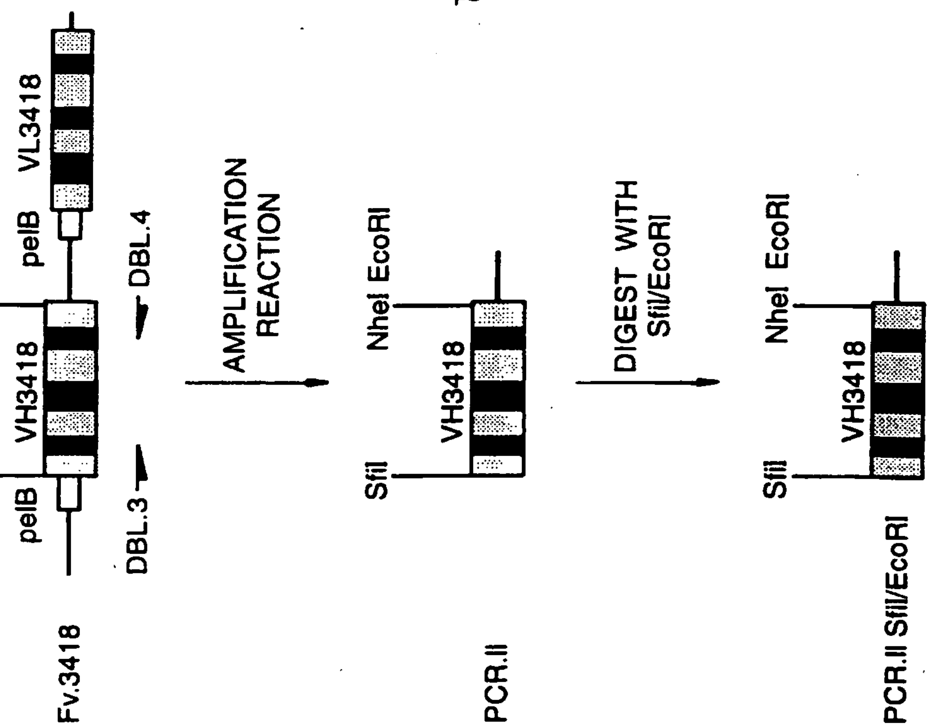

Fv.3418
pelB
PstI
VH3418
BstEII
pelB
VL3418
DBL.3
DBL.4
AMPLIFICATION REACTION
PCR.II
SfiI
VH3418
NheI EcoRI
DIGEST WITH SfiI/EcoRI
PCR.II SfiI/EcoRI
SfiI
VH3418
NheI EcoRI

Fig.20.

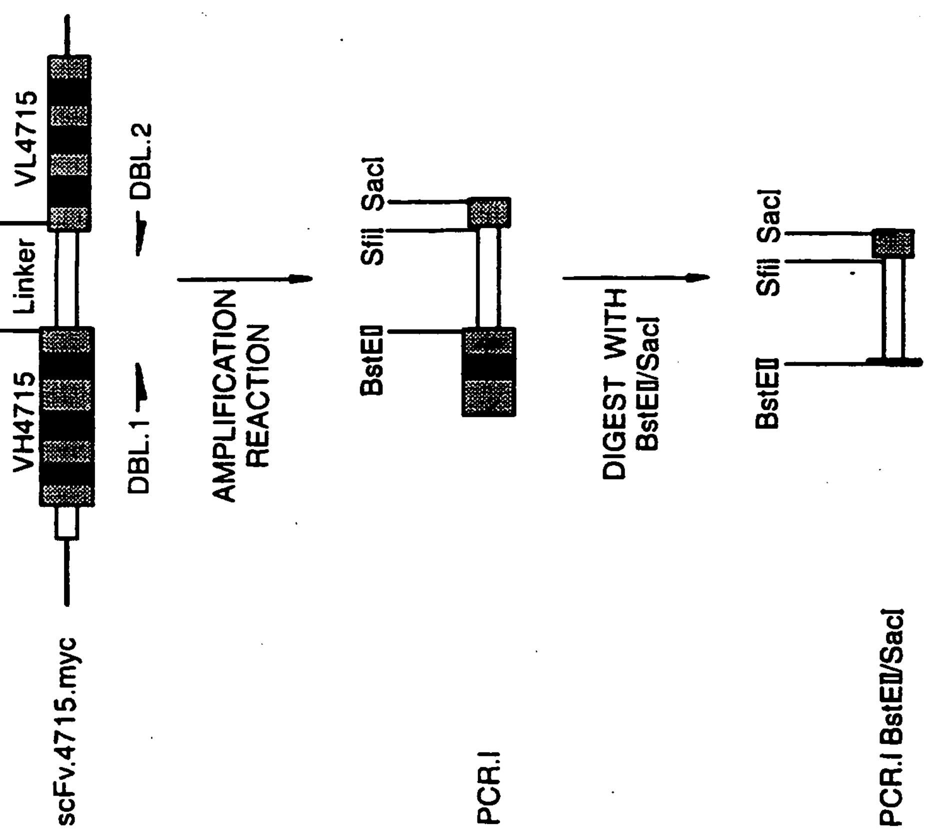

scFv.4715.myc
VH4715
BstEII
Linker
SacI
VL4715
DBL.1
DBL.2
AMPLIFICATION REACTION
PCR.I
BstEII
SfiI SacI
DIGEST WITH BstEII/SacI
PCR.I BstEII/SacI
BstEII
SfiI SacI

Fig.23.

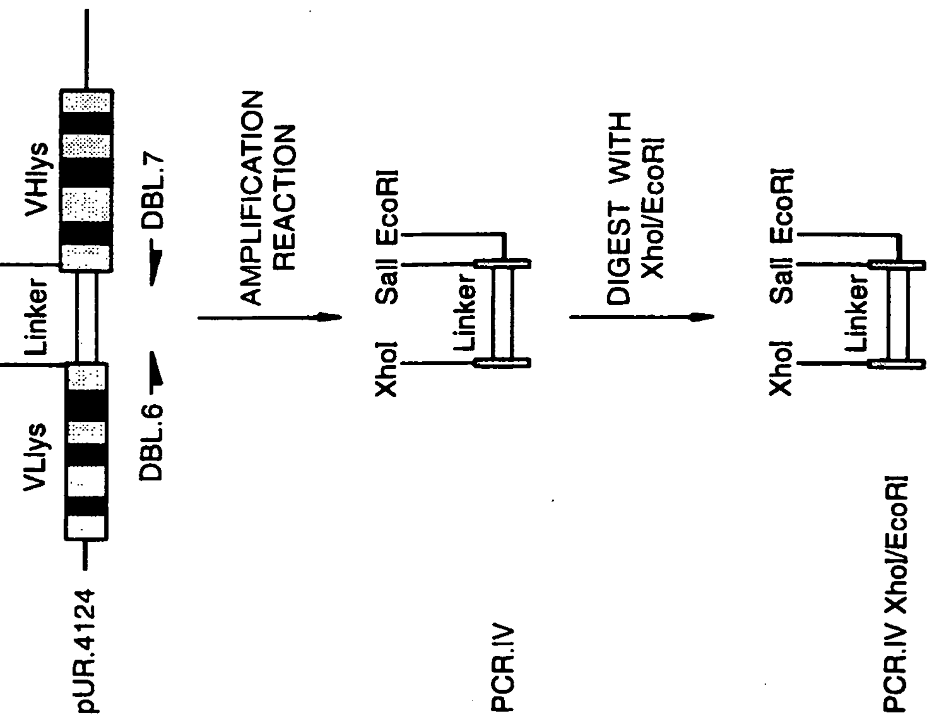

pUR.4124
VLlys
XhoI
Linker
PstI
VHlys
DBL.6
DBL.7
AMPLIFICATION REACTION
PCR.IV
XhoI
Linker
SalI
EcoRI
DIGEST WITH XhoI/EcoRI
PCR.IV XhoI/EcoRI
XhoI
Linker
SalI
EcoRI

Fig.22.

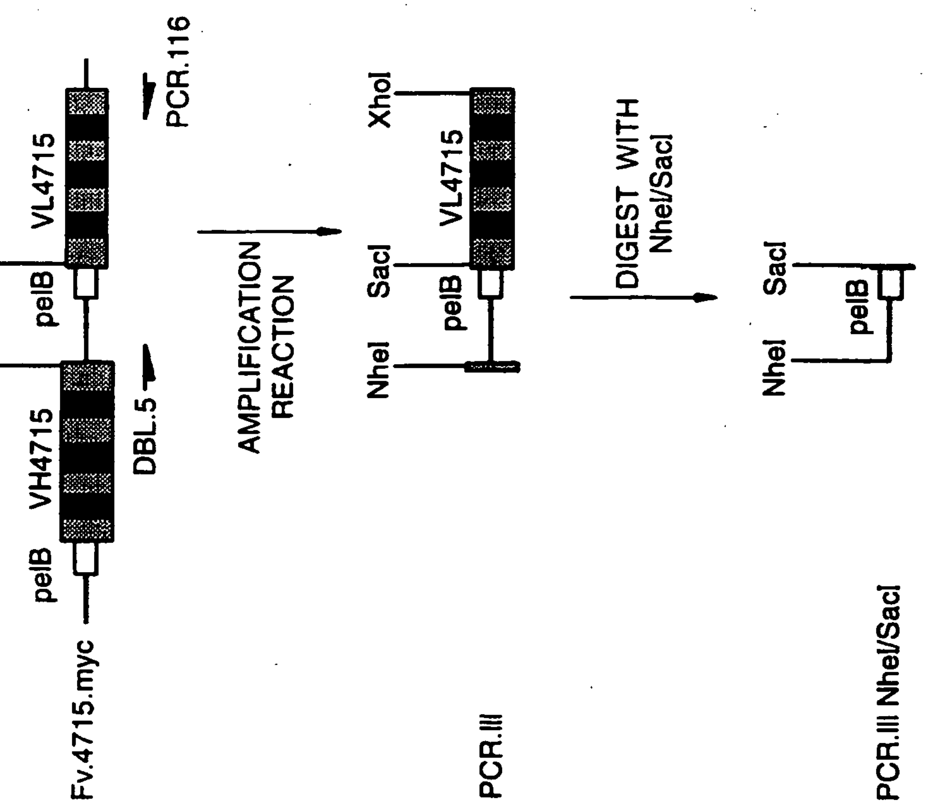

Fv.4715.myc
pelB
VH4715
BstEII
pelB
SacI
VL4715
DBL.5
PCR.116
AMPLIFICATION REACTION
PCR.III
NheI
pelB
SacI
VL4715
XhoI
DIGEST WITH NheI/SacI
PCR.III NheI/SacI
NheI
pelB
SacI

Fig.25.

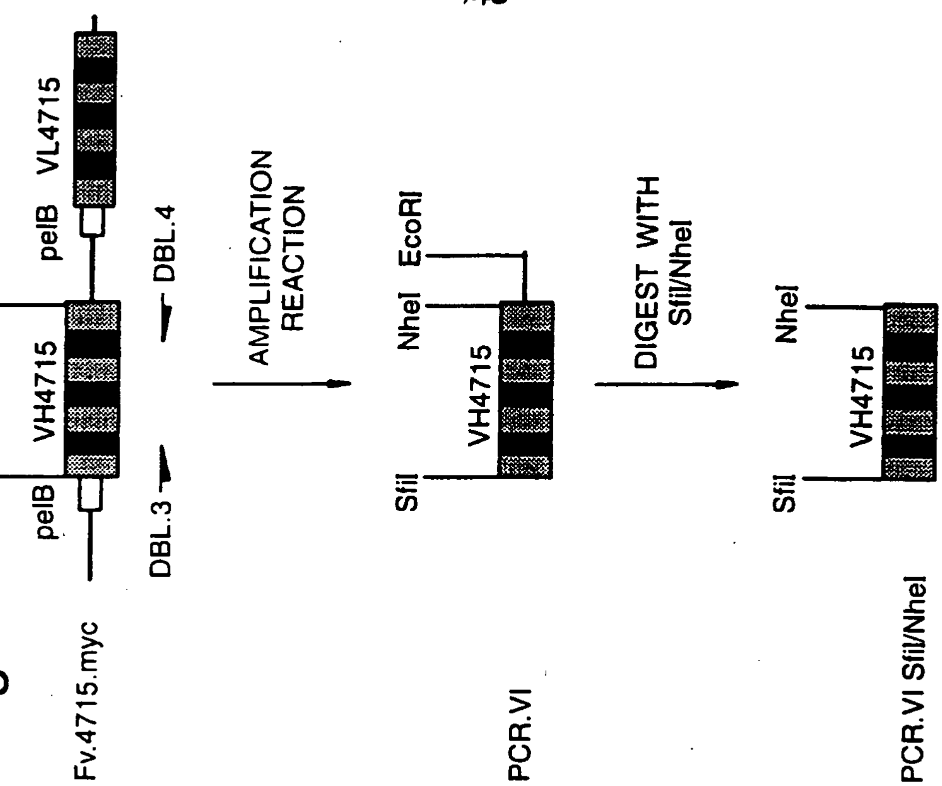

Fv.4715.myc
pelB
PstI
VH4715
BstEII
pelB
VL4715
DBL.3
DBL.4
AMPLIFICATION REACTION
SfiI
VH4715
NheI
EcoRI
PCR.VI
DIGEST WITH SfiI/NheI
SfiI
VH4715
NheI
PCR.VI SfiI/NheI

Fig.24.

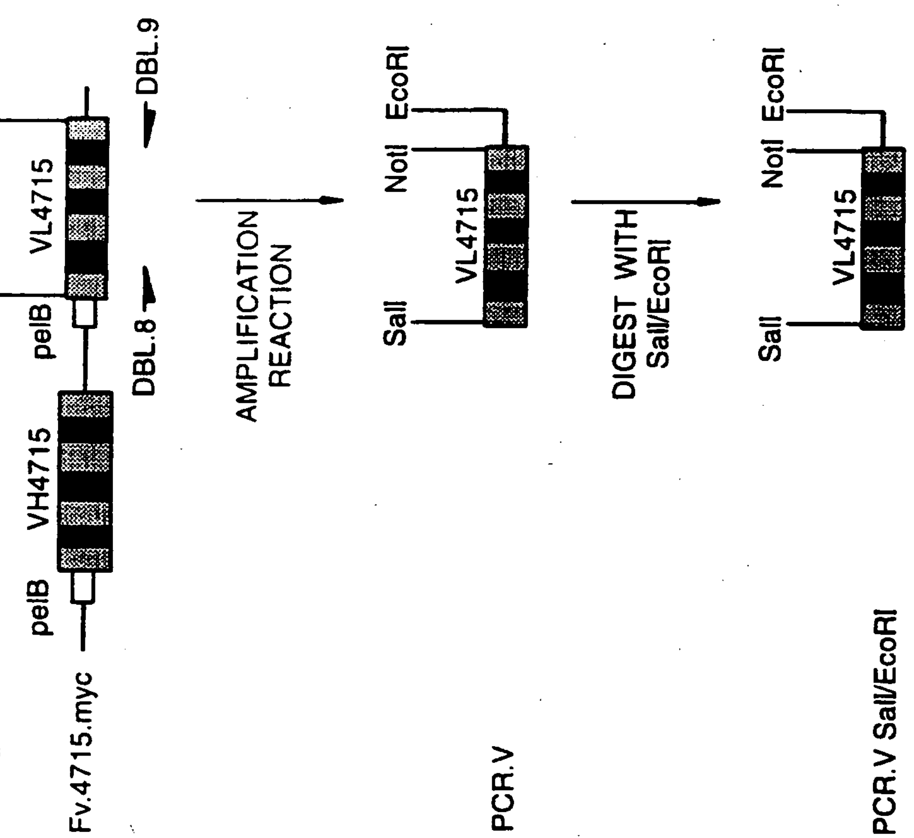

Fv.4715.myc
pelB
VH4715
pelB
SacI
VL4715
XhoI
DBL.8
DBL.9
AMPLIFICATION REACTION
SalI
VL4715
NotI
EcoRI
PCR.V
DIGEST WITH SalI/EcoRI
SalI
VL4715
NotI
EcoRI
PCR.V SalI/EcoRI

Fig.27.

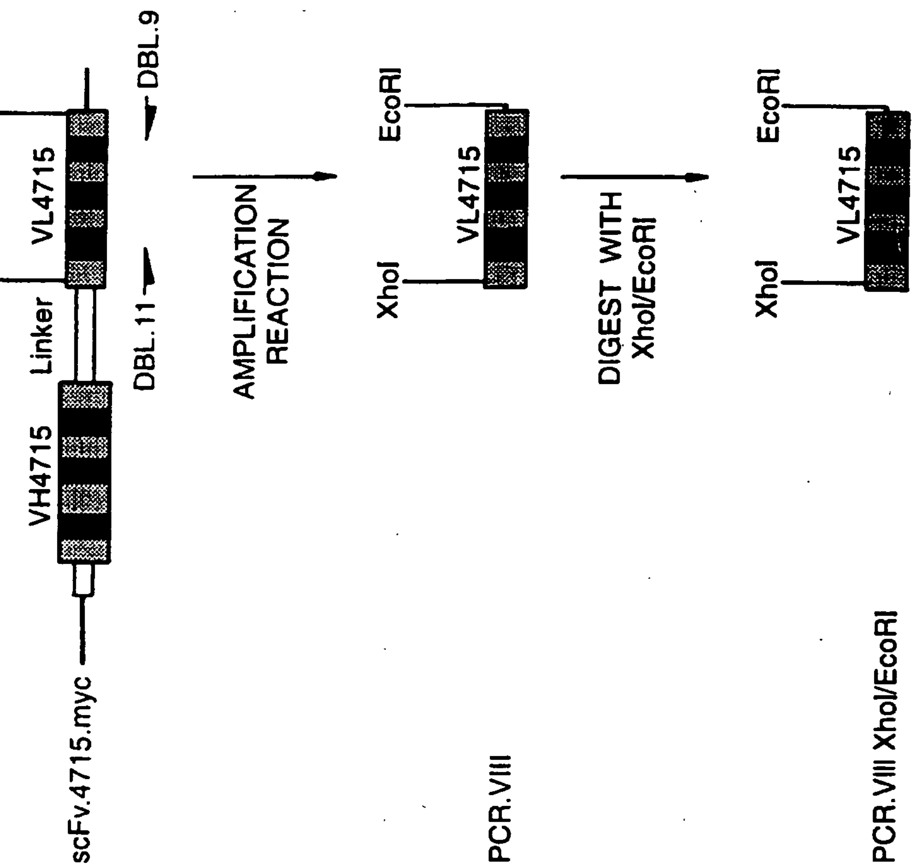

scFv.4715.myc
VH4715
Linker
SacI
VL4715
XhoI
DBL.11
DBL.9
AMPLIFICATION REACTION
PCR.VIII
XhoI
VL4715
EcoRI
DIGEST WITH XhoI/EcoRI
PCR.VIII XhoI/EcoRI
XhoI
VL4715
EcoRI

Fig.26.

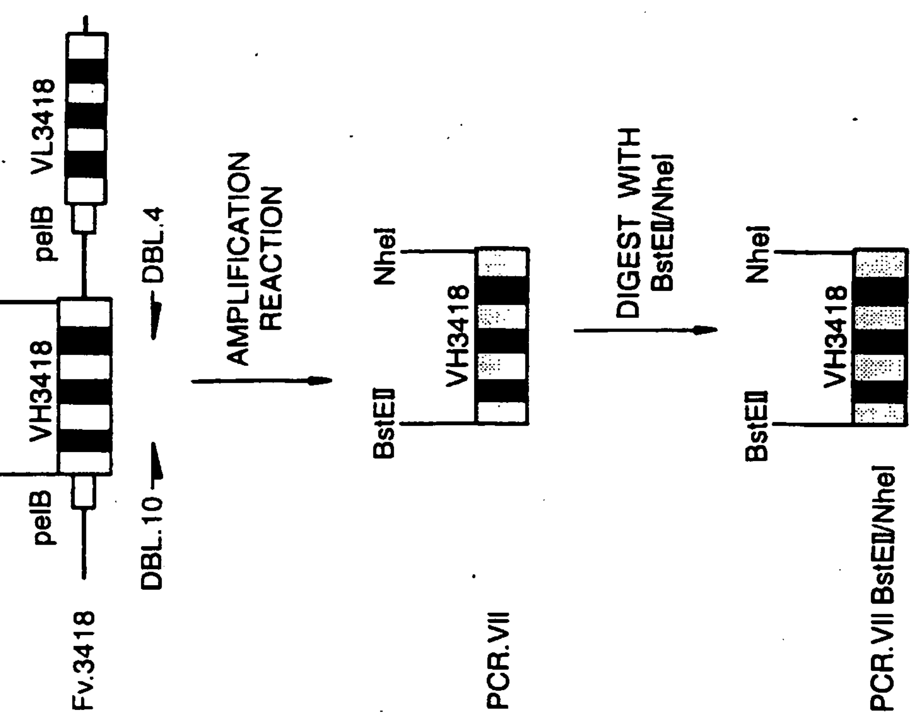

Fv.3418
pelB
PstI
VH3418
BstEII
pelB
VL3418
DBL.10
DBL.4
AMPLIFICATION REACTION
PCR.VII
BstEII
VH3418
NheI
DIGEST WITH BstEII/NheI
PCR.VII BstEII/NheI
BstEII
VH3418
NheI

Fig.28.

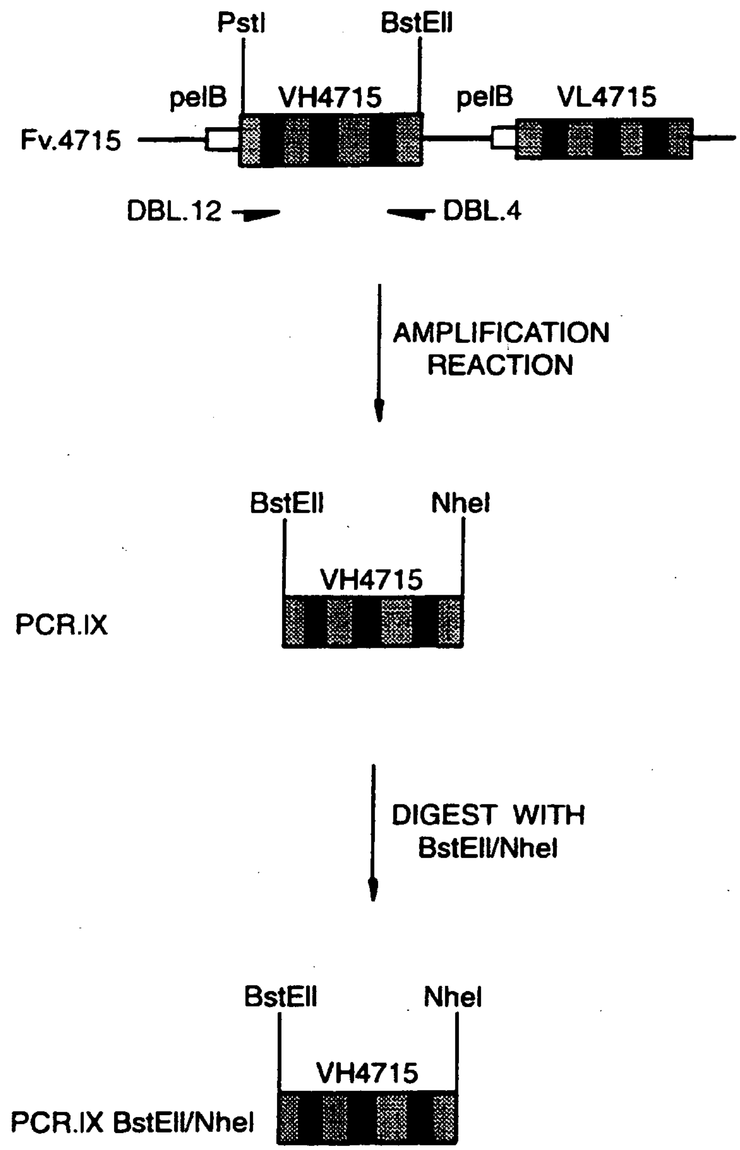

PstI
BstEII
pelB
VH4715
pelB
VL4715
Fv.4715
DBL.12
DBL.4
AMPLIFICATION
REACTION
BstEII
NheI
VH4715
PCR.IX
DIGEST WITH
BstEII/NheI
BstEII
NheI
VH4715
PCR.IX BstEII/NheI

Fig.29.

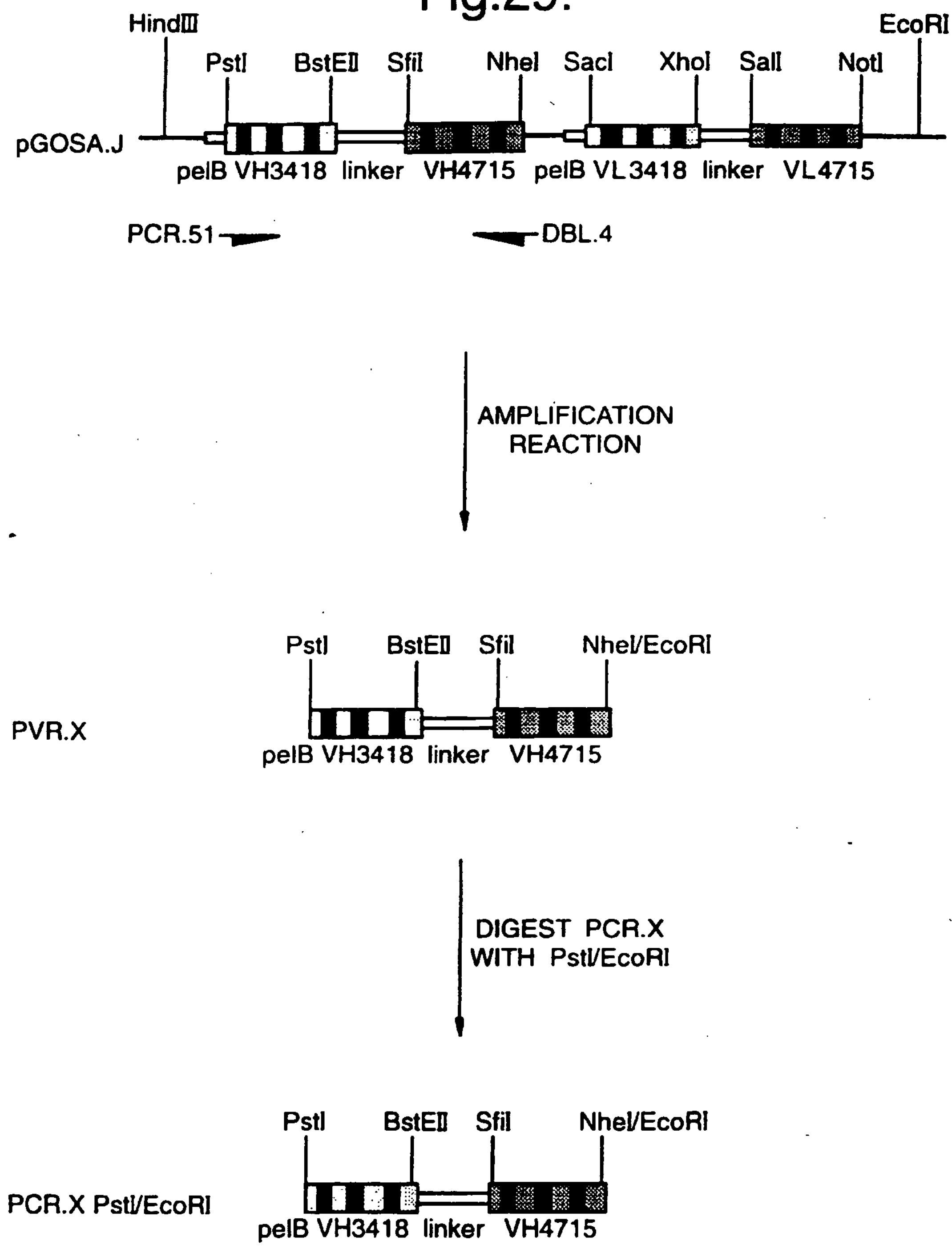

HindIII
EcoRI
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
pGOSA.J
pelB VH3418 linker VH4715 pelB VL3418 linker VL4715
PCR.51
DBL.4
AMPLIFICATION REACTION
PstI
BstEII
SfiI
NheI/EcoRI
PVR.X
pelB VH3418 linker VH4715
DIGEST PCR.X WITH PstI/EcoRI
PstI
BstEII
SfiI
NheI/EcoRI
PCR.X PstI/EcoRI
pelB VH3418 linker VH4715

Fig.30.

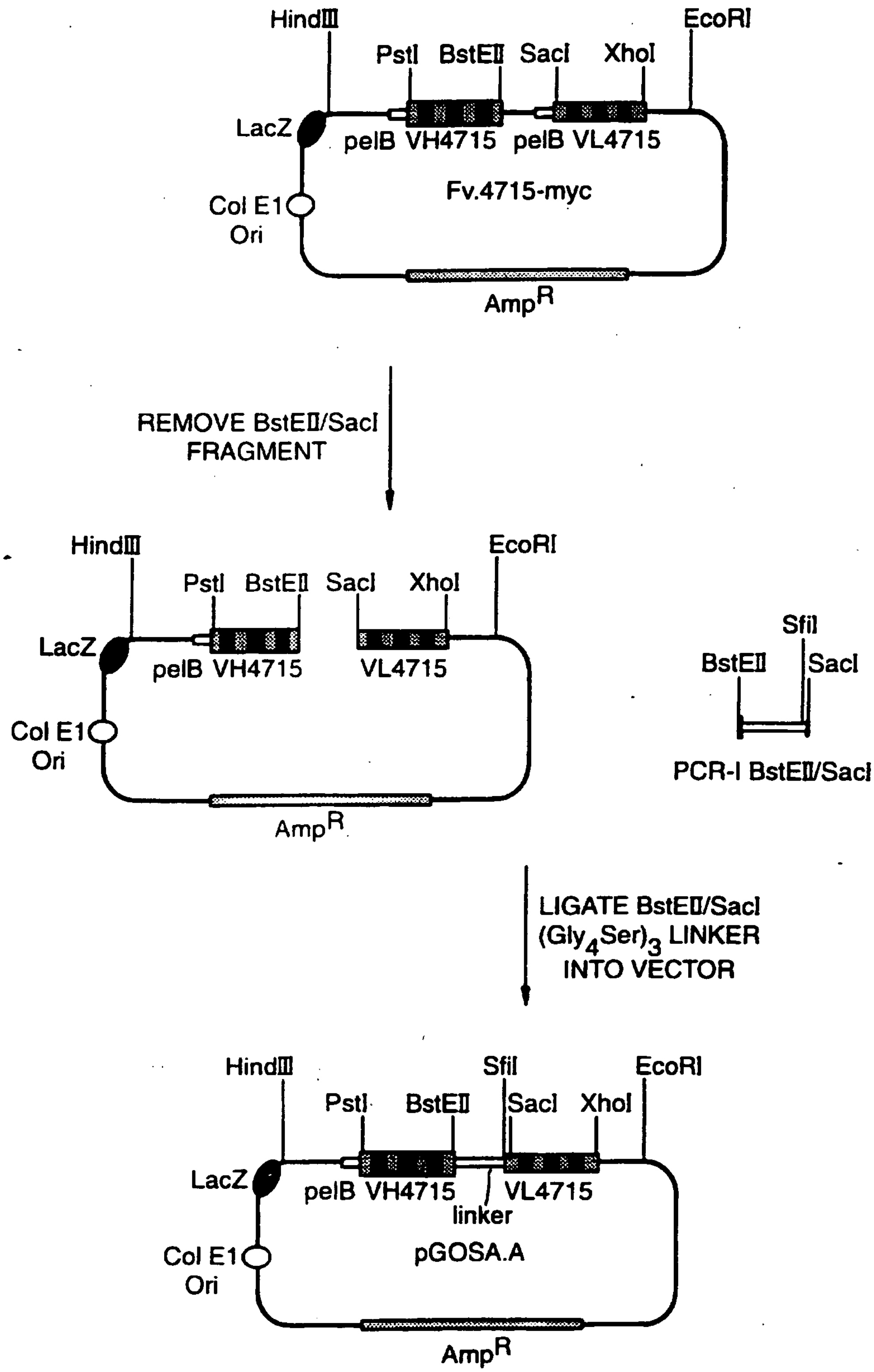

HindIII
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH4715
pelB VL4715
Fv.4715-myc
Col E1
Ori
AmpR
REMOVE BstEII/SacI FRAGMENT
HindIII
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH4715
VL4715
Col E1
Ori
AmpR
SfiI
BstEII
SacI
PCR-I BstEII/SacI
LIGATE BstEII/SacI
(Gly4Ser)3 LINKER
INTO VECTOR
HindIII
SfiI
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH4715
VL4715
linker
pGOSA.A
Col E1
Ori
AmpR

Fig.31.

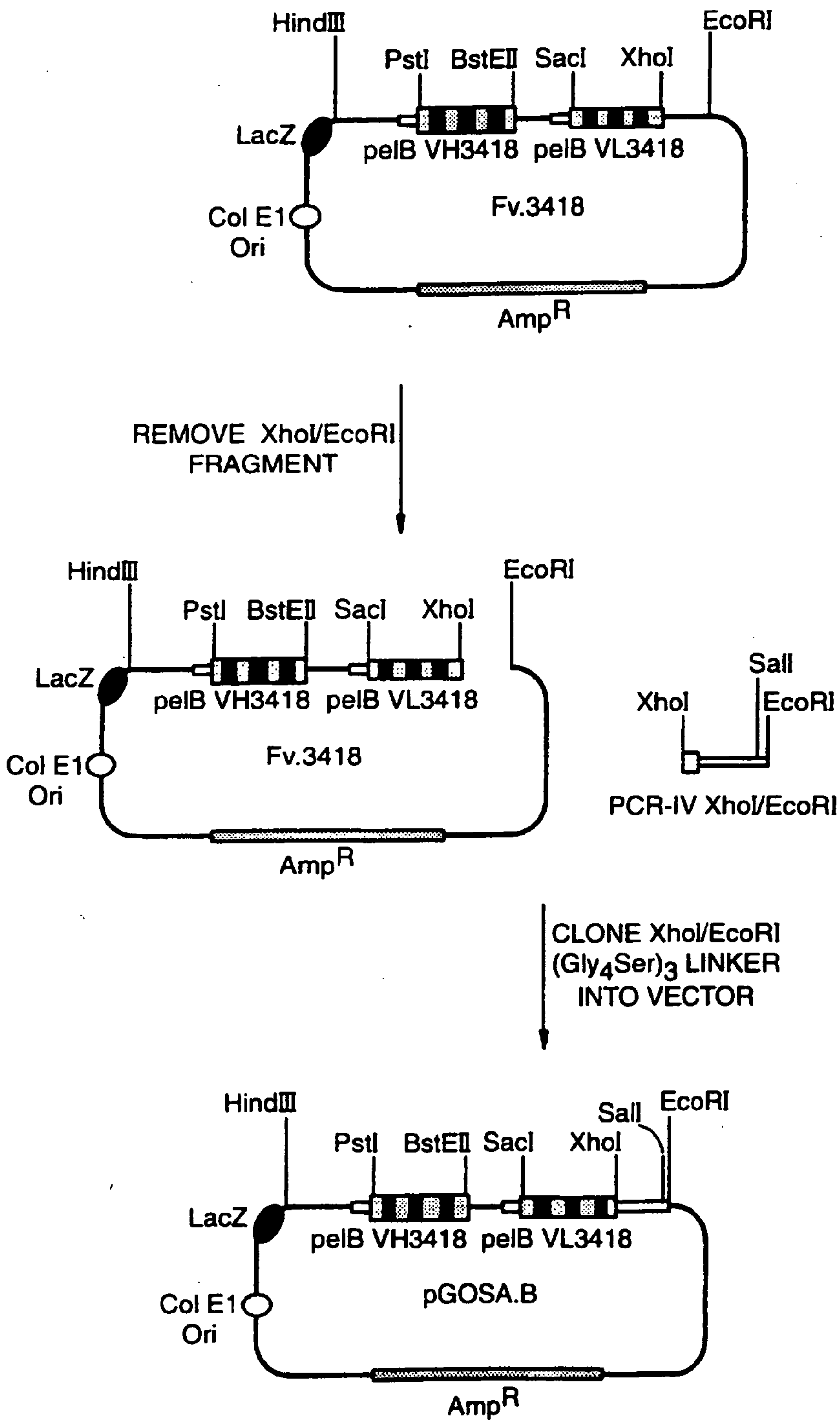

HindIII
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH3418
pelB VL3418
Fv.3418
Col E1
Ori
AmpR
REMOVE XhoI/EcoRI
FRAGMENT
SalI
PCR-IV XhoI/EcoRI
CLONE XhoI/EcoRI
(Gly4Ser)3 LINKER
INTO VECTOR
pGOSA.B

Fig.32.

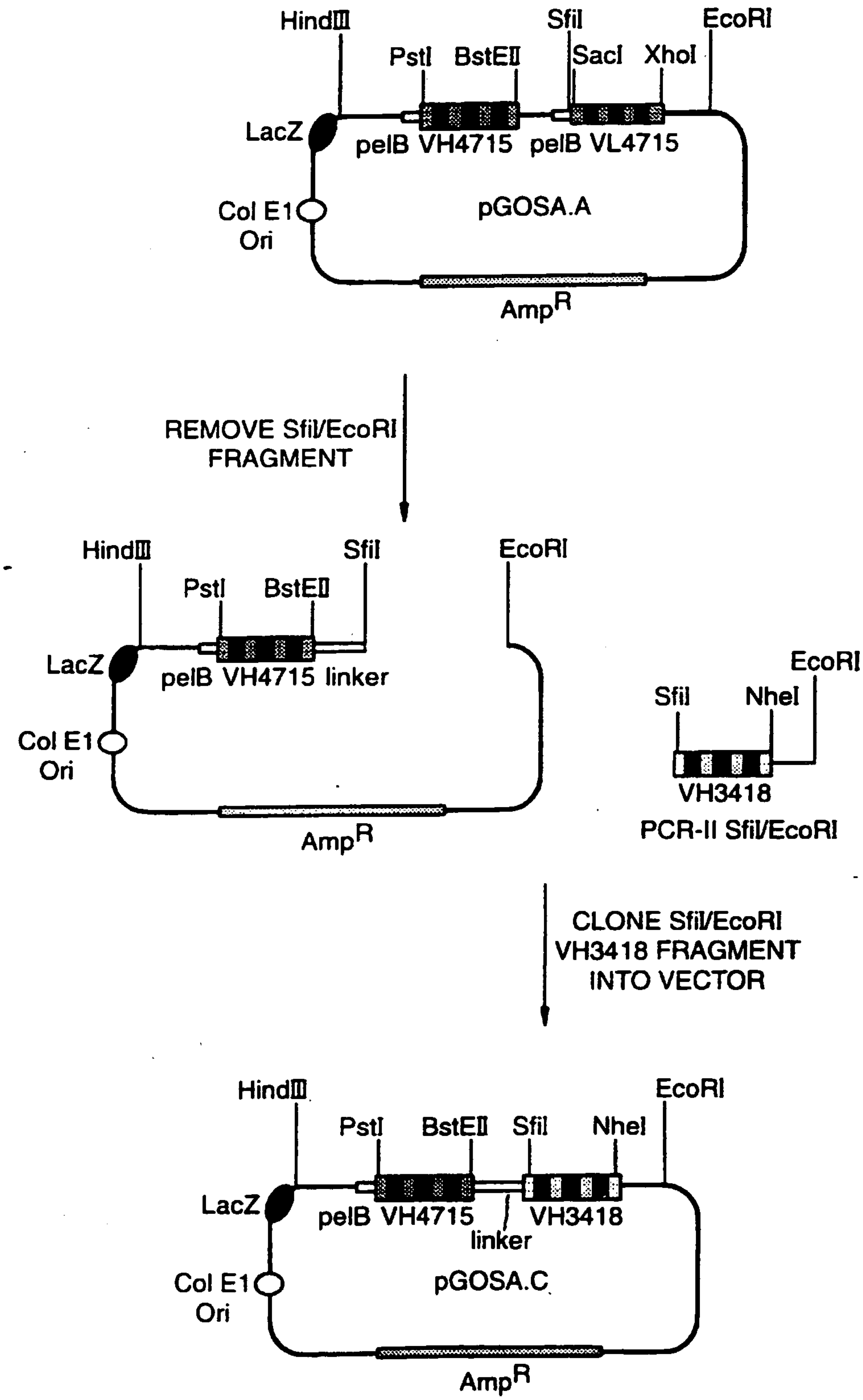

HindIII
SfiI
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH4715 pelB VL4715
Col E1
Ori
pGOSA.A
Amp^R
REMOVE SfiI/EcoRI
FRAGMENT
HindIII
SfiI
EcoRI
PstI
BstEII
LacZ
pelB VH4715 linker
Col E1
Ori
Amp^R
EcoRI
SfiI
NheI
VH3418
PCR-II SfiI/EcoRI
CLONE SfiI/EcoRI
VH3418 FRAGMENT
INTO VECTOR
HindIII
EcoRI
PstI
BstEII
SfiI
NheI
LacZ
pelB VH4715
VH3418
linker
Col E1
Ori
pGOSA.C
Amp^R

Fig.33.

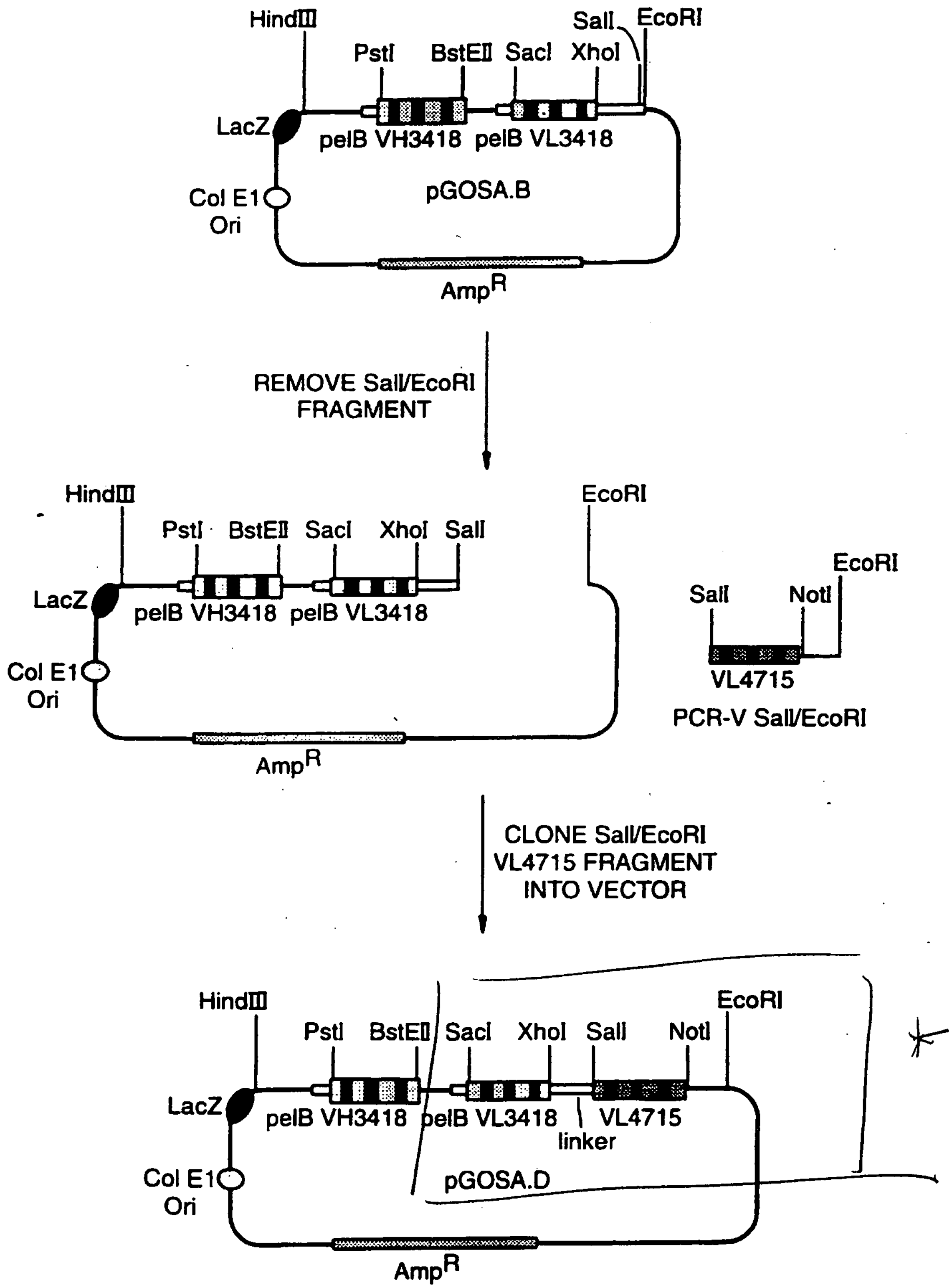

HindIII
PstI
BstEII
SacI
XhoI
SalI
EcoRI
LacZ
pelB VH3418
pelB VL3418
pGOSA.B
Col E1
Ori
AmpR
REMOVE SalI/EcoRI FRAGMENT
HindIII
PstI
BstEII
SacI
XhoI
SalI
EcoRI
LacZ
pelB VH3418
pelB VL3418
Col E1
Ori
AmpR
SalI
NotI
EcoRI
VL4715
PCR-V SalI/EcoRI
CLONE SalI/EcoRI VL4715 FRAGMENT INTO VECTOR
HindIII
PstI
BstEII
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
pelB VL3418
linker
VL4715
pGOSA.D
Col E1
Ori
AmpR

Fig.34.

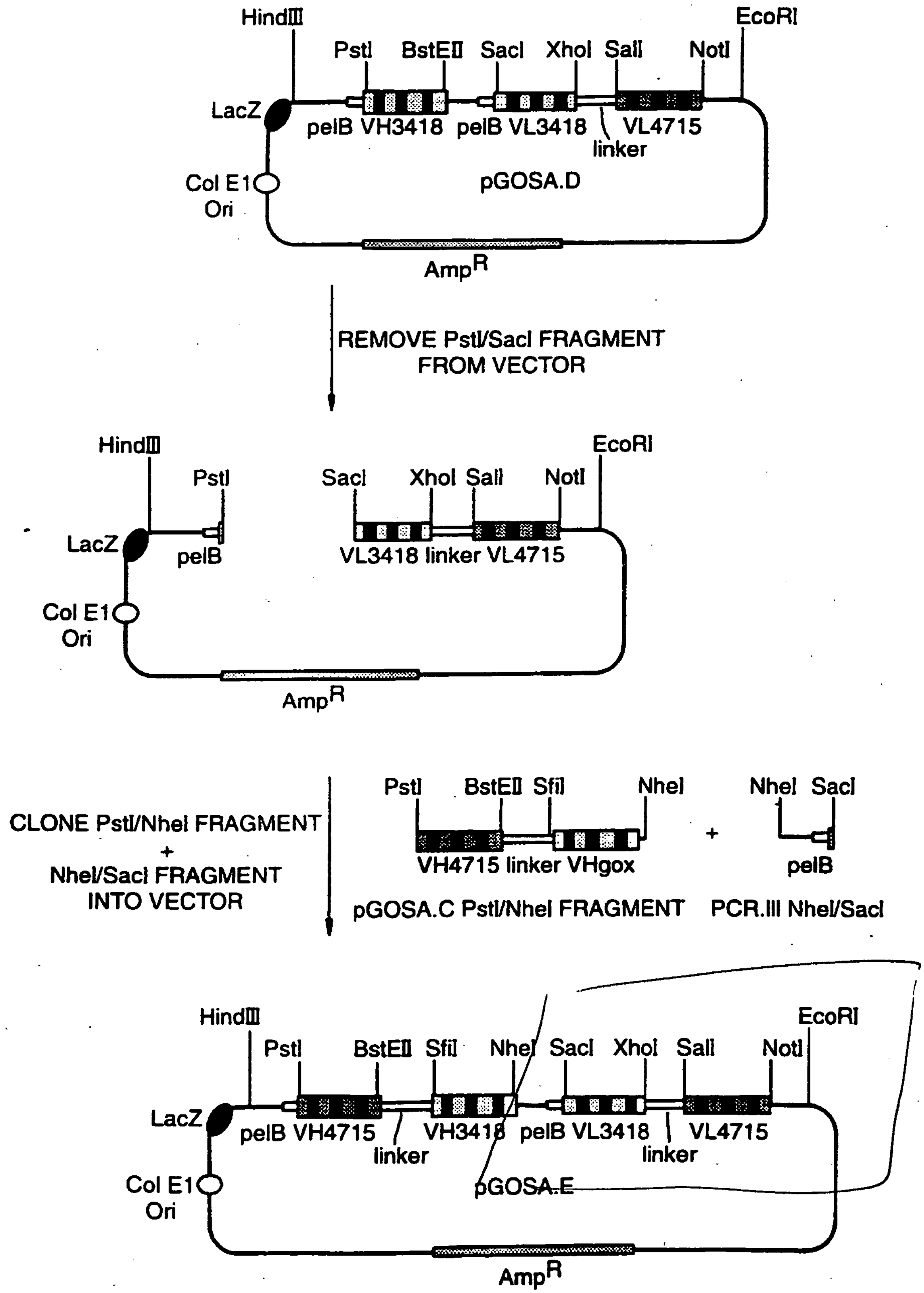

HindIII
PstI
BstEII
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
pelB VL3418
linker
VL4715
pGOSA.D
Col E1
Ori
AmpR
REMOVE PstI/SacI FRAGMENT
FROM VECTOR
HindIII
PstI
LacZ
pelB
SacI
XhoI
SalI
NotI
EcoRI
VL3418 linker VL4715
Col E1
Ori
AmpR
CLONE PstI/NheI FRAGMENT
+
NheI/SacI FRAGMENT
INTO VECTOR
PstI
BstEII
SfiI
NheI
VH4715 linker VHgox
pGOSA.C PstI/NheI FRAGMENT
+
NheI
SacI
pelB
PCR.III NheI/SacI
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH4715
linker
VH3418
pelB VL3418
linker
VL4715
pGOSA.E
Col E1
Ori
AmpR

Fig.35.

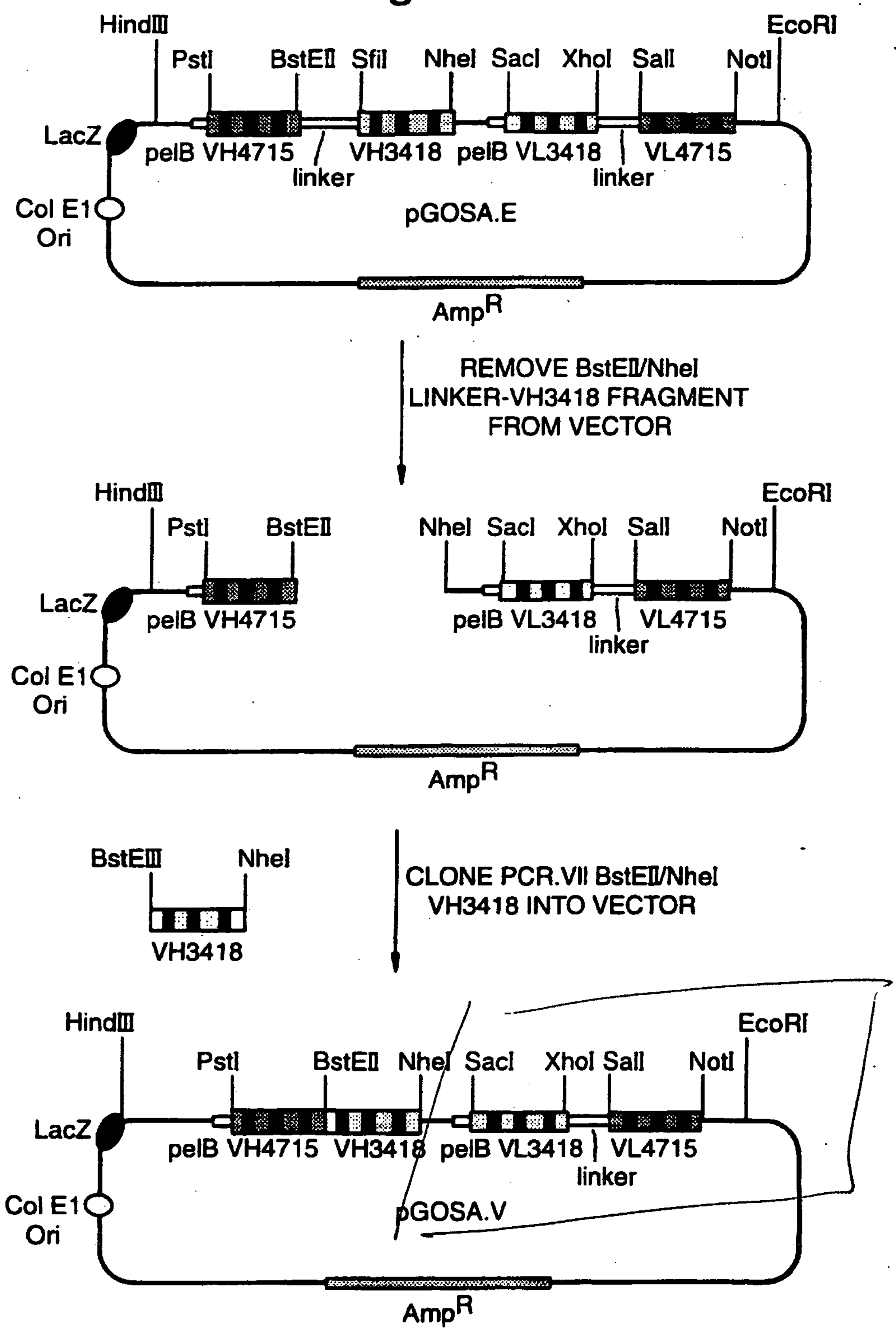

HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH4715
linker
VH3418
pelB VL3418
linker
VL4715
Col E1 Ori
pGOSA.E
AmpR
REMOVE BstEII/NheI
LINKER-VH3418 FRAGMENT
FROM VECTOR
CLONE PCR.VII BstEII/NheI
VH3418 INTO VECTOR
VH3418
pGOSA.V

Fig.36.

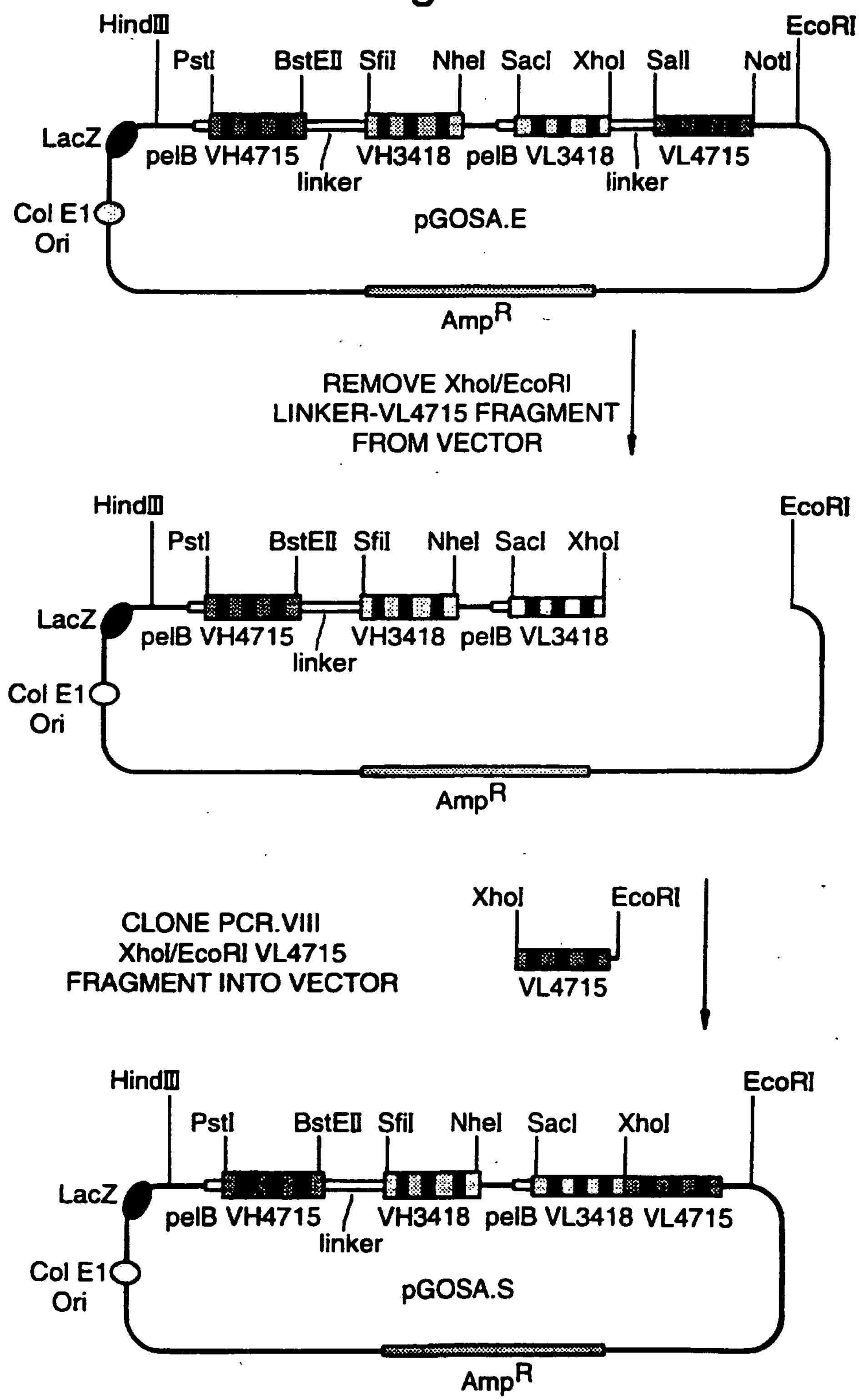

HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH4715
linker
VH3418
pelB VL3418
linker
VL4715
Col E1 Ori
pGOSA.E
AmpR
REMOVE XhoI/EcoRI
LINKER-VL4715 FRAGMENT
FROM VECTOR
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
EcoRI
LacZ
pelB VH4715
linker
VH3418
pelB VL3418
Col E1 Ori
AmpR
CLONE PCR.VIII
XhoI/EcoRI VL4715
FRAGMENT INTO VECTOR
XhoI
EcoRI
VL4715
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
EcoRI
LacZ
pelB VH4715
linker
VH3418
pelB VL3418
VL4715
Col E1 Ori
pGOSA.S
AmpR

Fig.37.

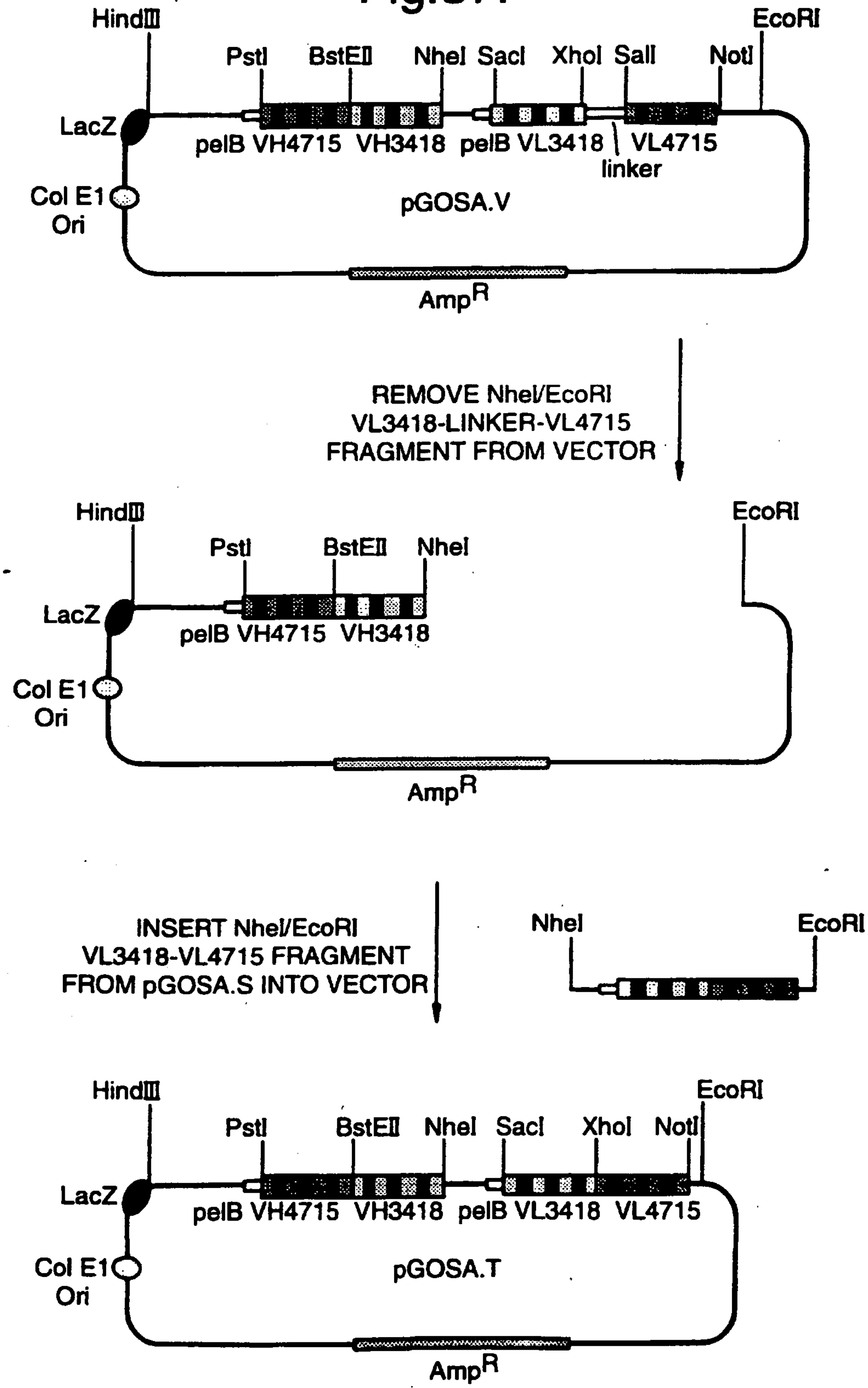

HindIII
PstI
BstEII
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH4715 VH3418 pelB VL3418 VL4715
linker
Col E1 Ori
pGOSA.V
AmpR
REMOVE NheI/EcoRI
VL3418-LINKER-VL4715
FRAGMENT FROM VECTOR
HindIII
PstI
BstEII
NheI
EcoRI
LacZ
pelB VH4715 VH3418
Col E1 Ori
AmpR
INSERT NheI/EcoRI
VL3418-VL4715 FRAGMENT
FROM pGOSA.S INTO VECTOR
NheI
EcoRI
HindIII
PstI
BstEII
NheI
SacI
XhoI
NotI
EcoRI
LacZ
pelB VH4715 VH3418 pelB VL3418 VL4715
Col E1 Ori
pGOSA.T
AmpR

Fig.38(1/2).

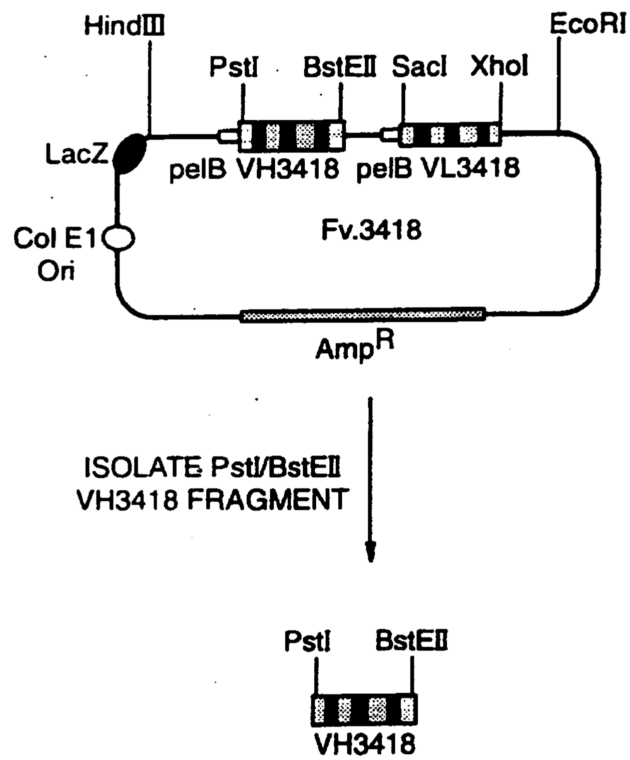

HindIII
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH3418
pelB VL3418
Col E1
Ori
Fv.3418
AmpR
ISOLATE PstI/BstEII
VH3418 FRAGMENT
PstI
BstEII
VH3418

Fig.38(2/2).

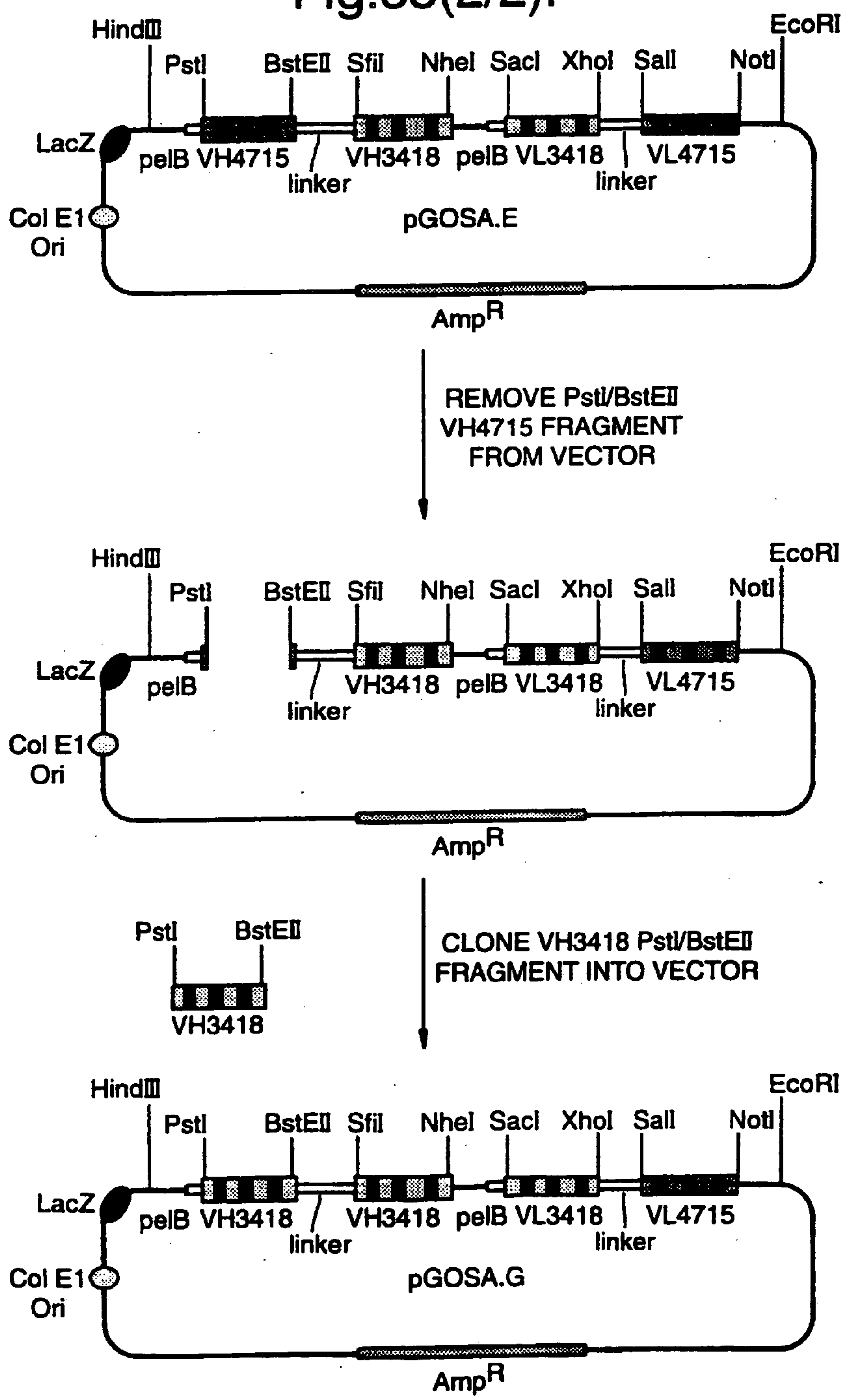

HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH4715
linker
VH3418
pelB VL3418
linker
VL4715
Col E1
Ori
pGOSA.E
AmpR
REMOVE PstI/BstEII
VH4715 FRAGMENT
FROM VECTOR
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB
linker
VH3418
pelB VL3418
linker
VL4715
Col E1
Ori
AmpR
PstI
BstEII
VH3418
CLONE VH3418 PstI/BstEII
FRAGMENT INTO VECTOR
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
linker
VH3418
pelB VL3418
linker
VL4715
Col E1
Ori
pGOSA.G
AmpR

Fig.39.

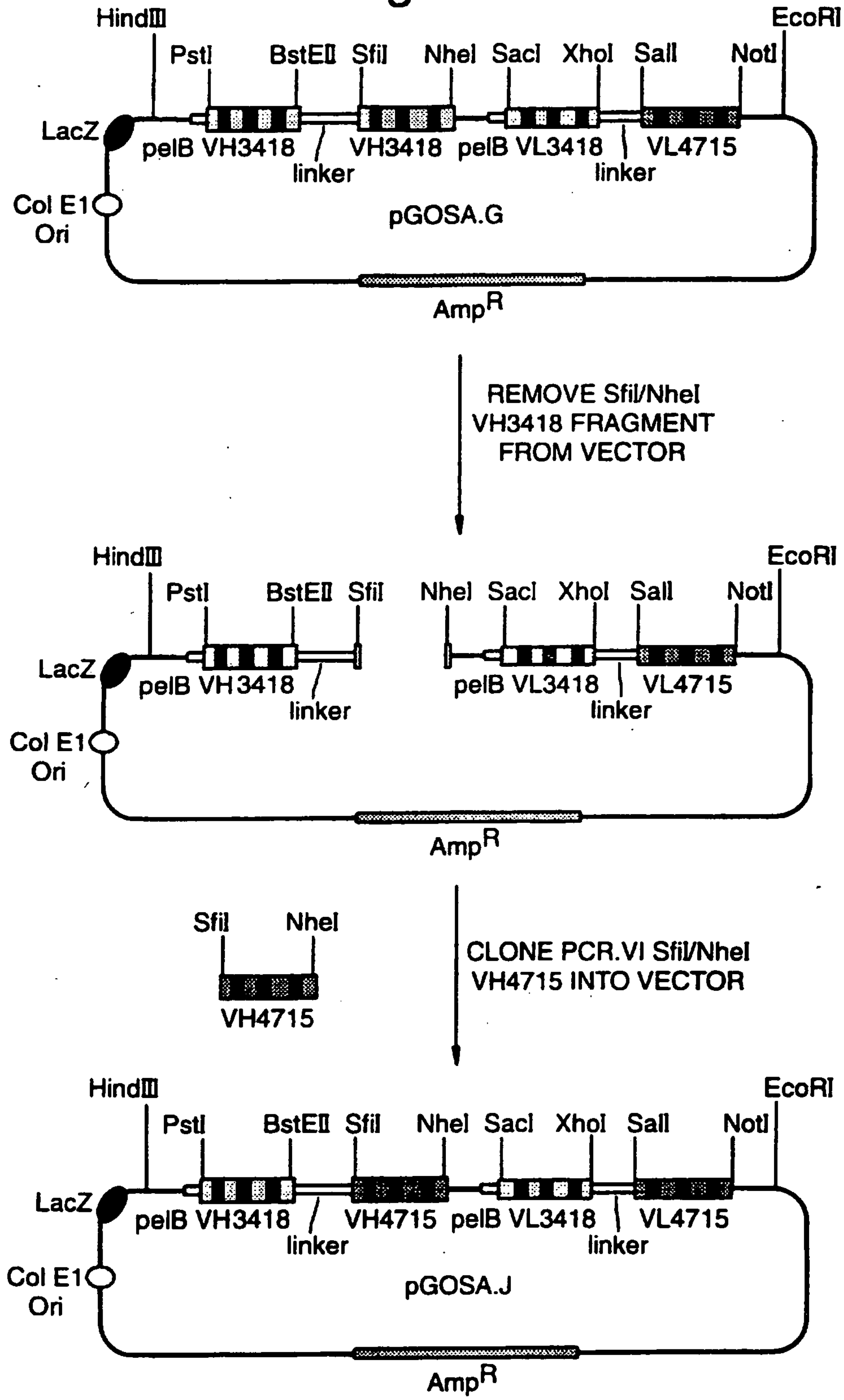

HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
linker
VH3418
pelB VL3418
linker
VL4715
Col E1
Ori
pGOSA.G
AmpR
REMOVE SfiI/NheI
VH3418 FRAGMENT
FROM VECTOR
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
linker
pelB VL3418
linker
VL4715
Col E1
Ori
AmpR
SfiI
NheI
VH4715
CLONE PCR.VI SfiI/NheI
VH4715 INTO VECTOR
HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
linker
VH4715
pelB VL3418
linker
VL4715
Col E1
Ori
pGOSA.J
AmpR

Fig.40.

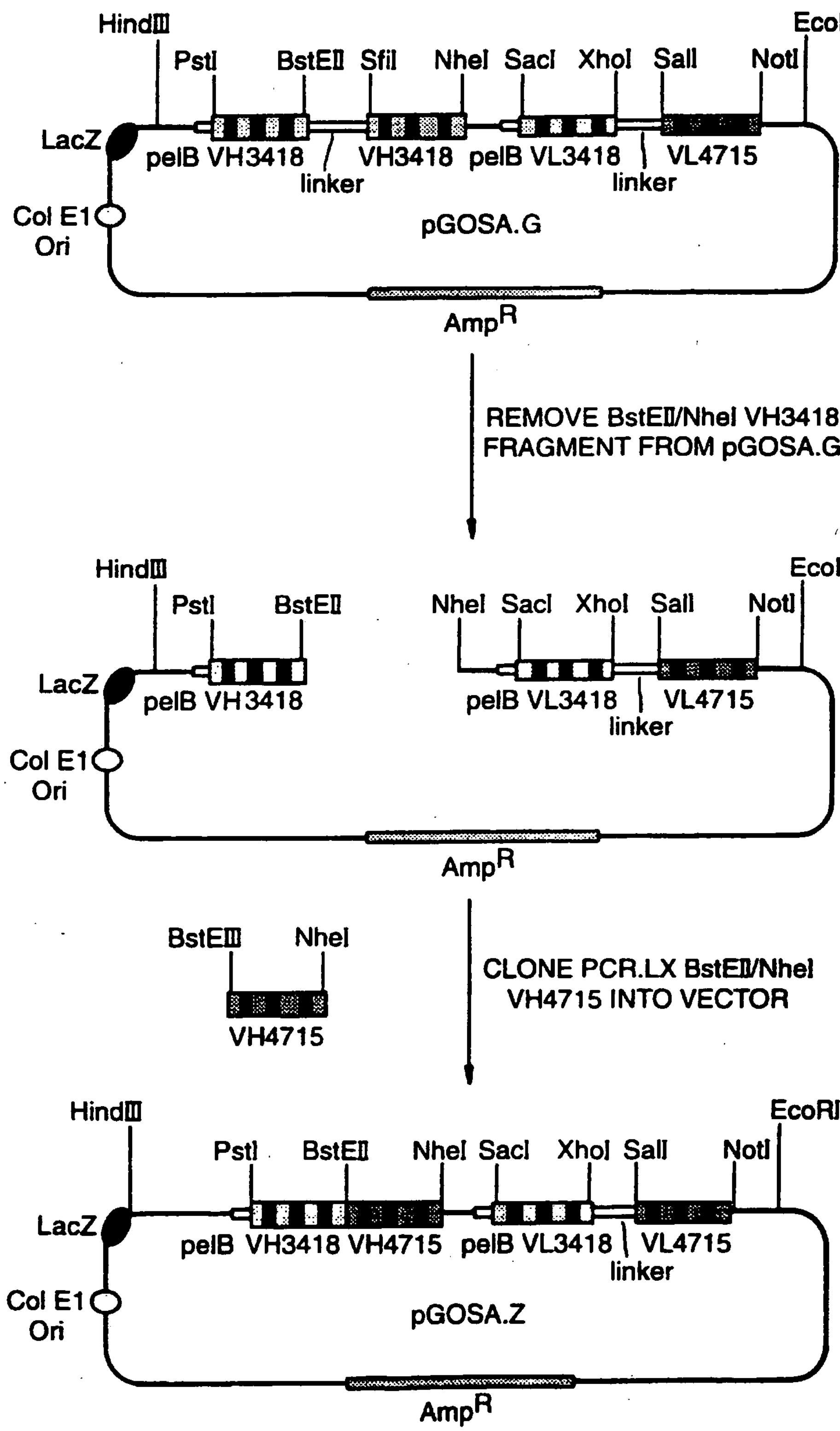

HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
linker
VH3418
pelB VL3418
linker
VL4715
Col E1 Ori
pGOSA.G
AmpR
REMOVE BstEII/NheI VH3418 FRAGMENT FROM pGOSA.G
HindIII
PstI
BstEII
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
pelB VL3418
linker
VL4715
Col E1 Ori
AmpR
BstEII
NheI
VH4715
CLONE PCR.LX BstEII/NheI VH4715 INTO VECTOR
HindIII
PstI
BstEII
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418 VH4715
pelB VL3418
linker
VL4715
Col E1 Ori
pGOSA.Z
AmpR

Fig.41.

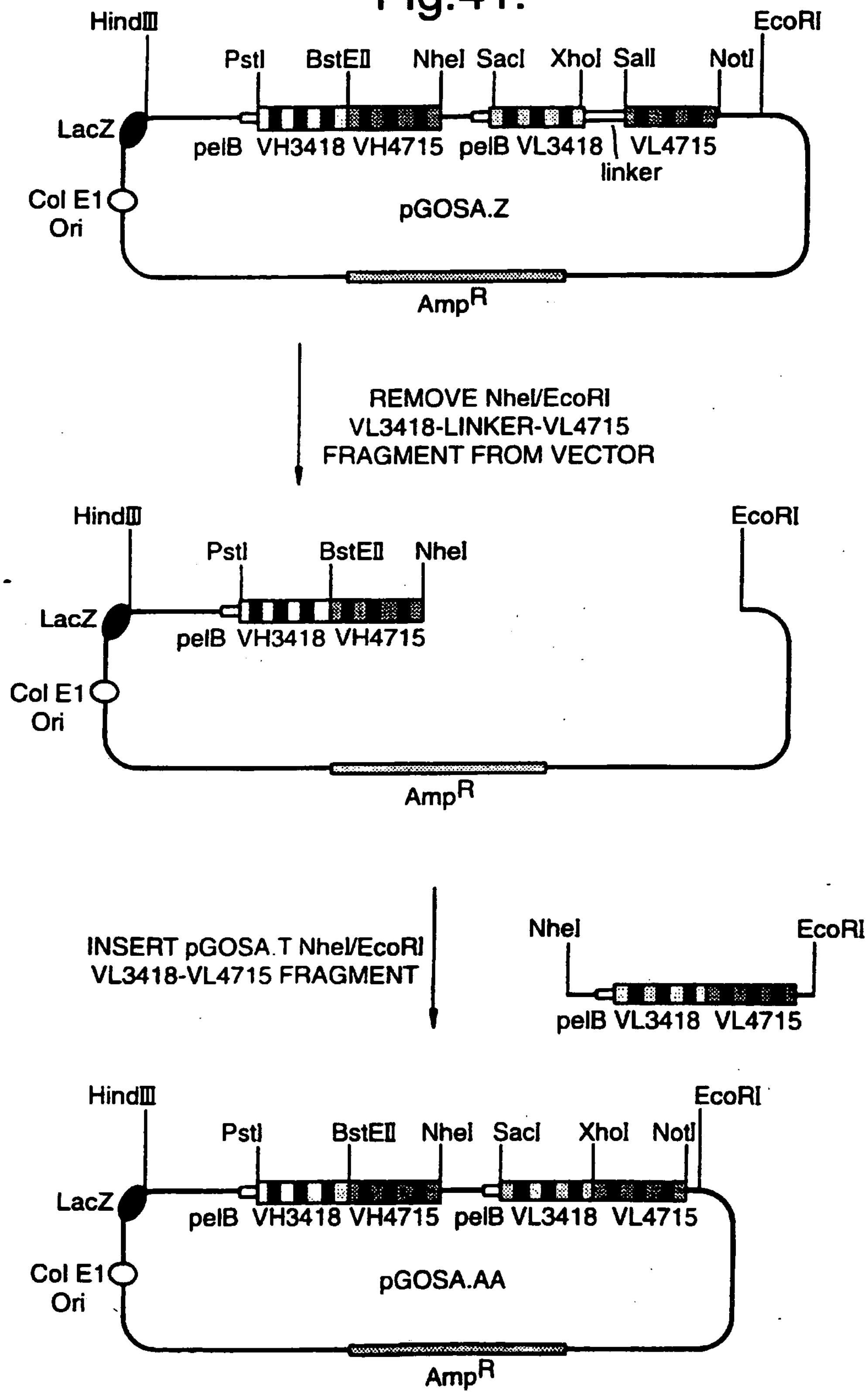

HindIII
PstI
BstEII
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418 VH4715 pelB VL3418
VL4715
linker
Col E1
Ori
pGOSA.Z
AmpR
REMOVE NheI/EcoRI
VL3418-LINKER-VL4715
FRAGMENT FROM VECTOR
HindIII
PstI
BstEII
NheI
EcoRI
LacZ
pelB VH3418 VH4715
Col E1
Ori
AmpR
INSERT pGOSA.T NheI/EcoRI
VL3418-VL4715 FRAGMENT
NheI
EcoRI
pelB VL3418 VL4715
HindIII
PstI
BstEII
NheI
SacI
XhoI
NotI
EcoRI
LacZ
pelB VH3418 VH4715 pelB VL3418 VL4715
Col E1
Ori
pGOSA.AA
AmpR

Fig.42.

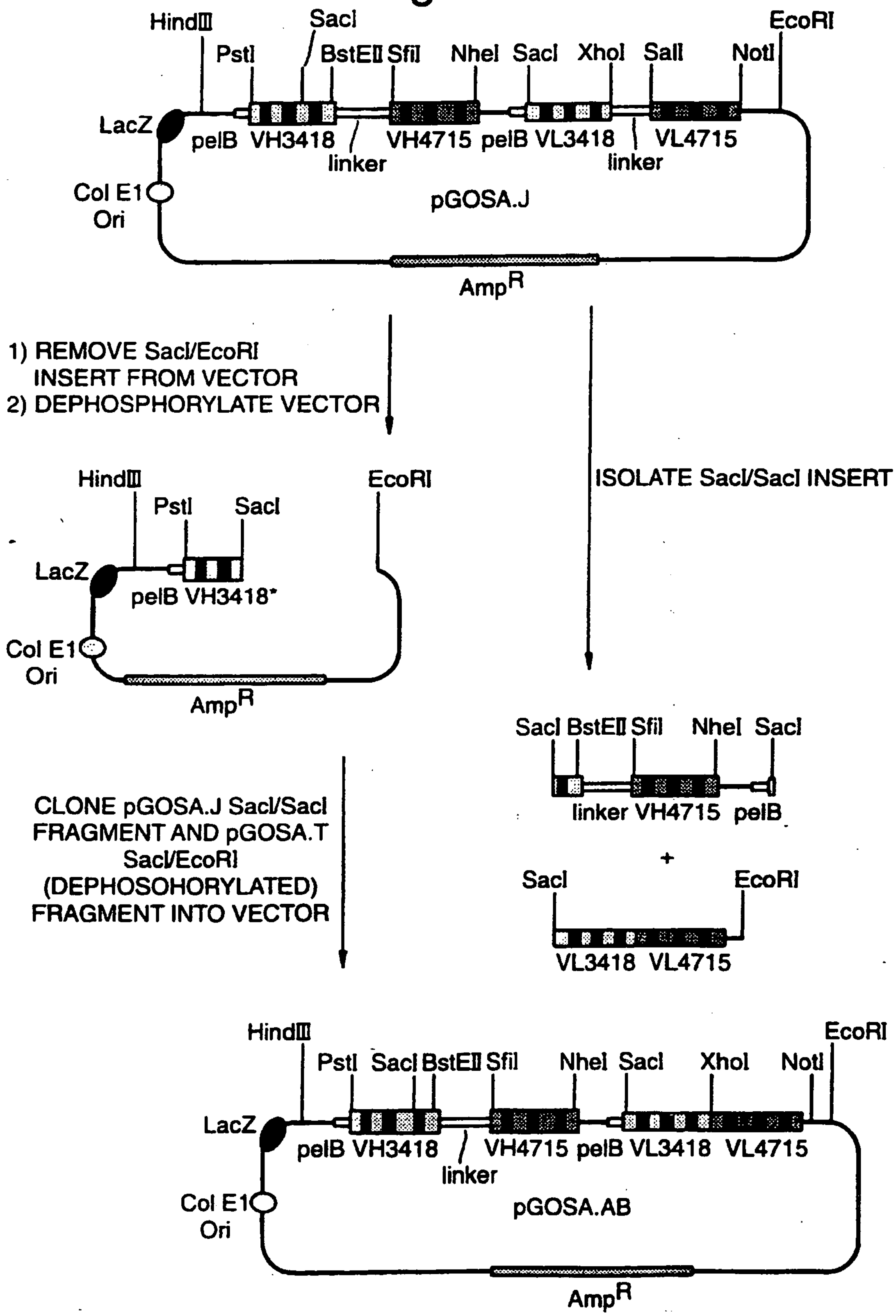

HindIII
SacI
PstI
BstEII SfiI
NheI SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418
linker
VH4715 pelB VL3418
linker
VL4715
Col E1
Ori
pGOSA.J
AmpR
1) REMOVE SacI/EcoRI
INSERT FROM VECTOR
2) DEPHOSPHORYLATE VECTOR
ISOLATE SacI/SacI INSERT
HindIII
PstI
SacI
EcoRI
LacZ
pelB VH3418*
Col E1
Ori
AmpR
SacI BstEII SfiI
NheI SacI
linker VH4715 pelB
+
SacI
EcoRI
VL3418 VL4715
CLONE pGOSA.J SacI/SacI
FRAGMENT AND pGOSA.T
SacI/EcoRI
(DEPHOSOHORYLATED)
FRAGMENT INTO VECTOR
HindIII
PstI SacI BstEII SfiI
NheI SacI
XhoI
NotI
EcoRI
LacZ
pelB VH3418
linker
VH4715 pelB VL3418 VL4715
Col E1
Ori
pGOSA.AB
AmpR

Fig.43.

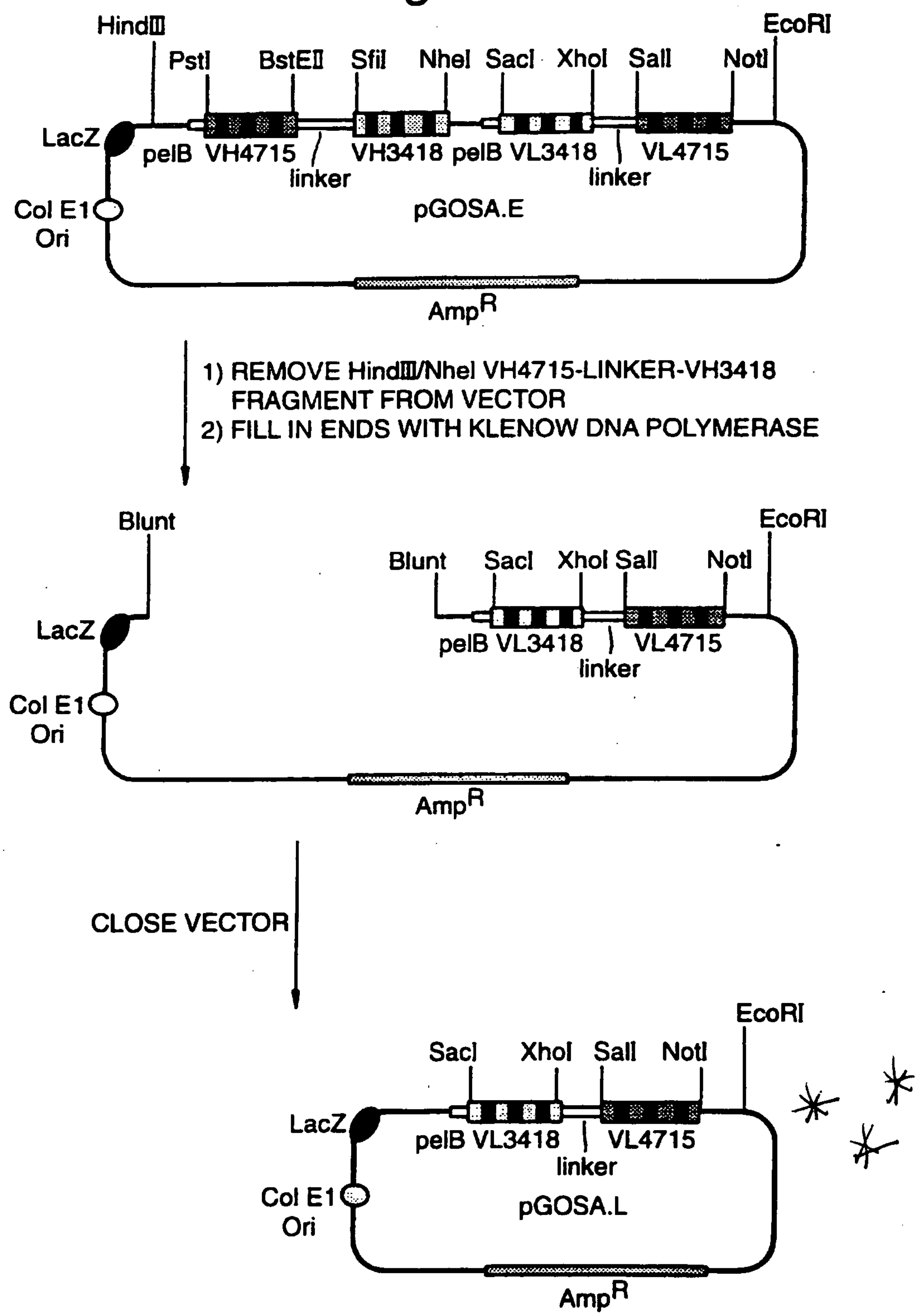

HindIII
PstI
BstEII
SfiI
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB
VH4715
linker
VH3418
pelB
VL3418
linker
VL4715
Col E1
Ori
pGOSA.E
AmpR
1) REMOVE HindIII/NheI VH4715-LINKER-VH3418 FRAGMENT FROM VECTOR
2) FILL IN ENDS WITH KLENOW DNA POLYMERASE
Blunt
Blunt
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VL3418
linker
VL4715
Col E1
Ori
AmpR
CLOSE VECTOR
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VL3418
linker
VL4715
Col E1
Ori
pGOSA.L
AmpR

Fig.44.

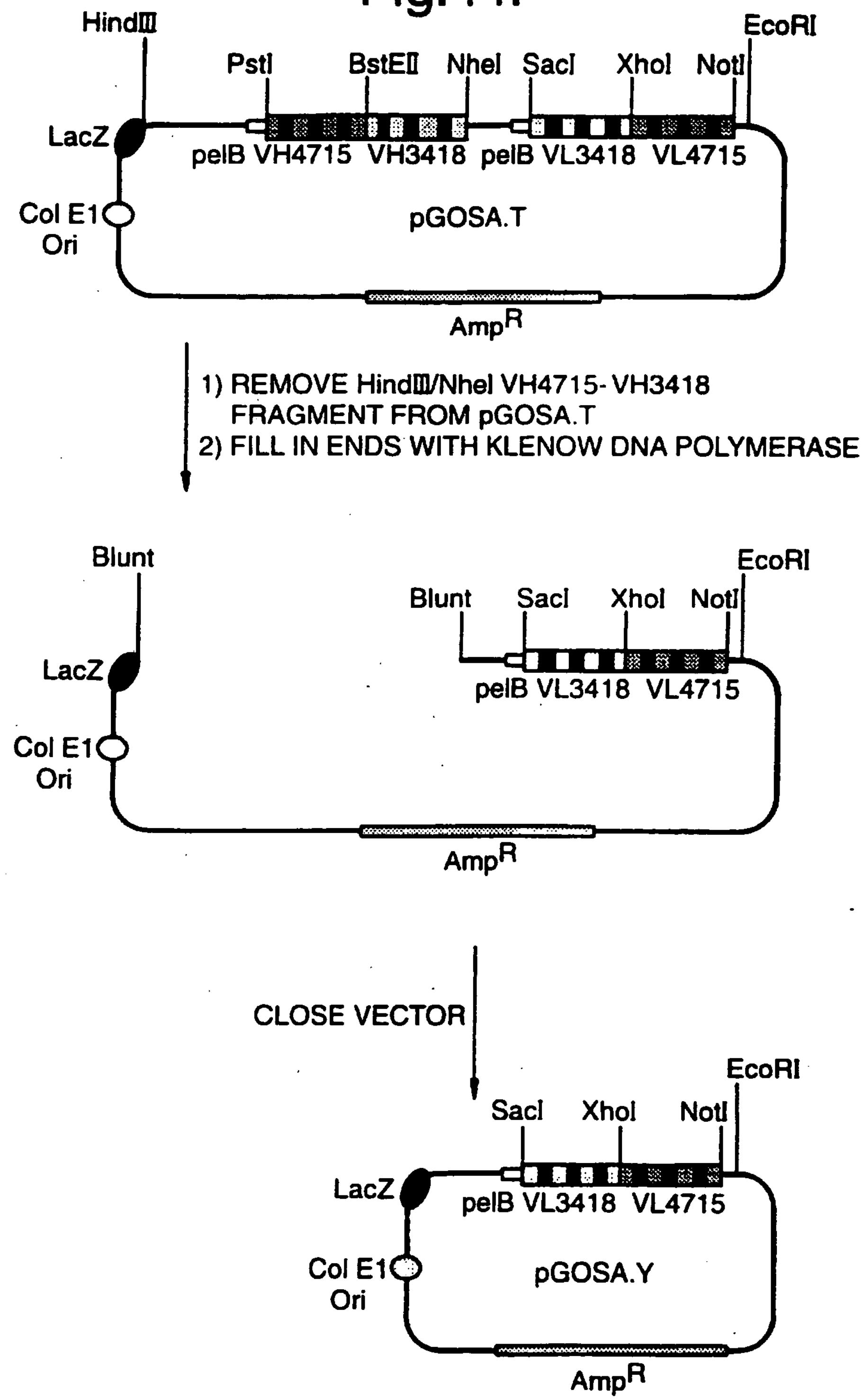

HindIII
PstI
BstEII
NheI
SacI
XhoI
NotI
EcoRI
LacZ
pelB VH4715 VH3418 pelB VL3418 VL4715
Col E1
Ori
pGOSA.T
AmpR
1) REMOVE HindIII/NheI VH4715-VH3418
FRAGMENT FROM pGOSA.T
2) FILL IN ENDS WITH KLENOW DNA POLYMERASE
Blunt
Blunt
SacI
XhoI
NotI
EcoRI
LacZ
pelB VL3418 VL4715
Col E1
Ori
AmpR
CLOSE VECTOR
SacI
XhoI
NotI
EcoRI
LacZ
pelB VL3418 VL4715
Col E1
Ori
pGOSA.Y
AmpR

Fig.45.

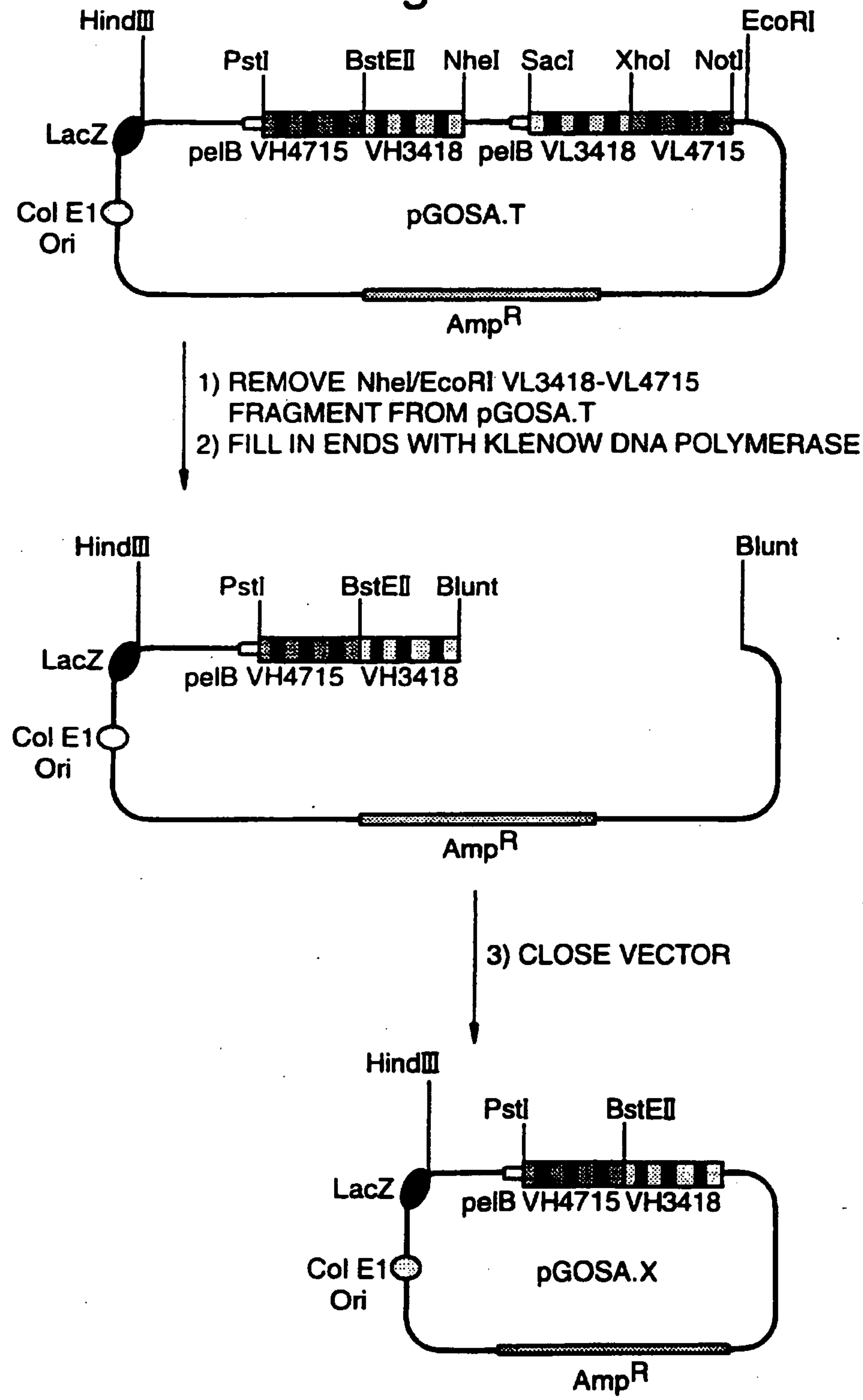

HindIII
PstI
BstEII
NheI
SacI
XhoI
NotI
EcoRI
LacZ
pelB VH4715 VH3418 pelB VL3418 VL4715
Col E1
Ori
pGOSA.T
AmpR
1) REMOVE NheI/EcoRI VL3418-VL4715
FRAGMENT FROM pGOSA.T
2) FILL IN ENDS WITH KLENOW DNA POLYMERASE
HindIII
PstI
BstEII
Blunt
Blunt
LacZ
pelB VH4715 VH3418
Col E1
Ori
AmpR
3) CLOSE VECTOR
HindIII
PstI
BstEII
LacZ
pelB VH4715 VH3418
Col E1
Ori
pGOSA.X
AmpR

Fig.46.

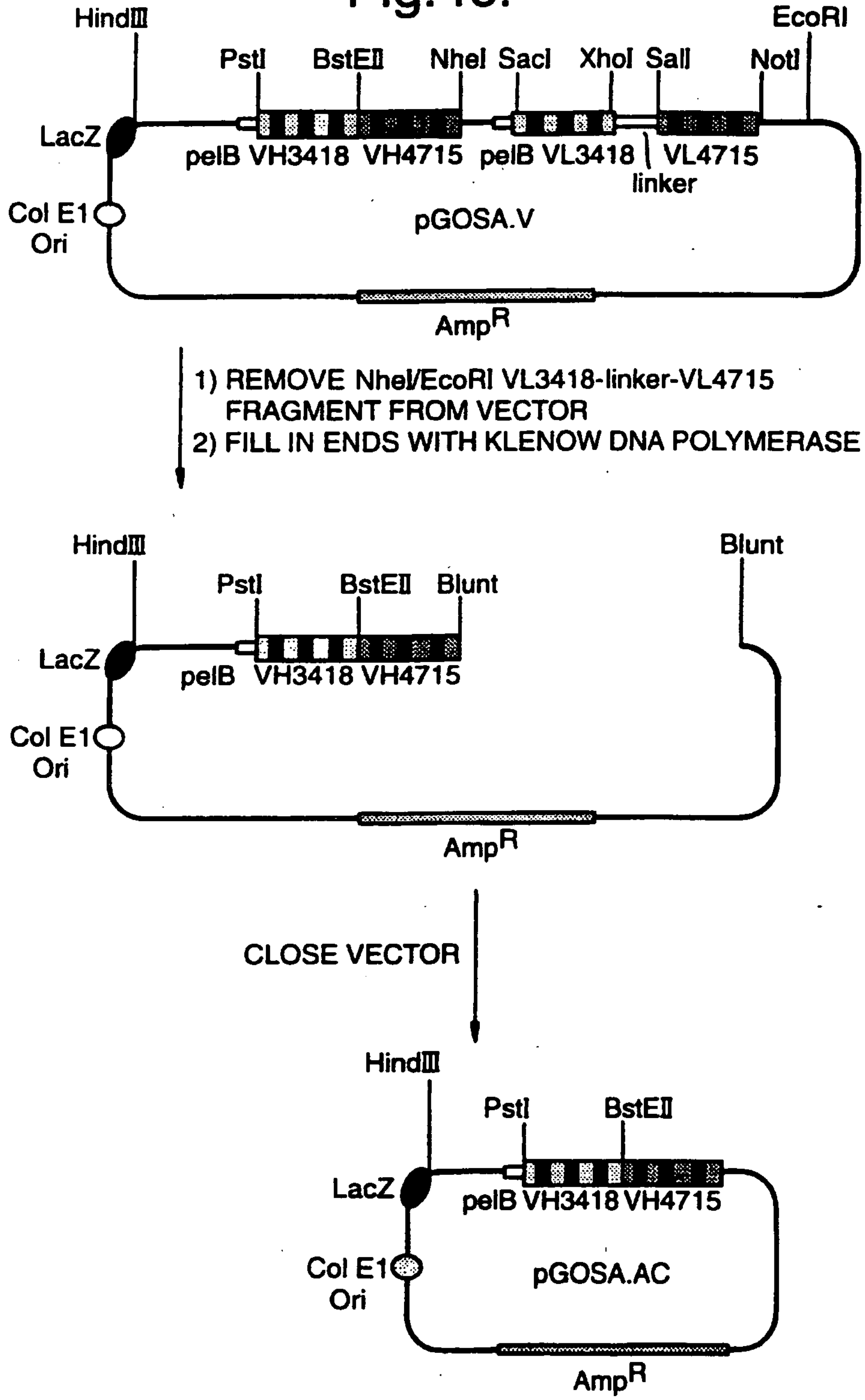
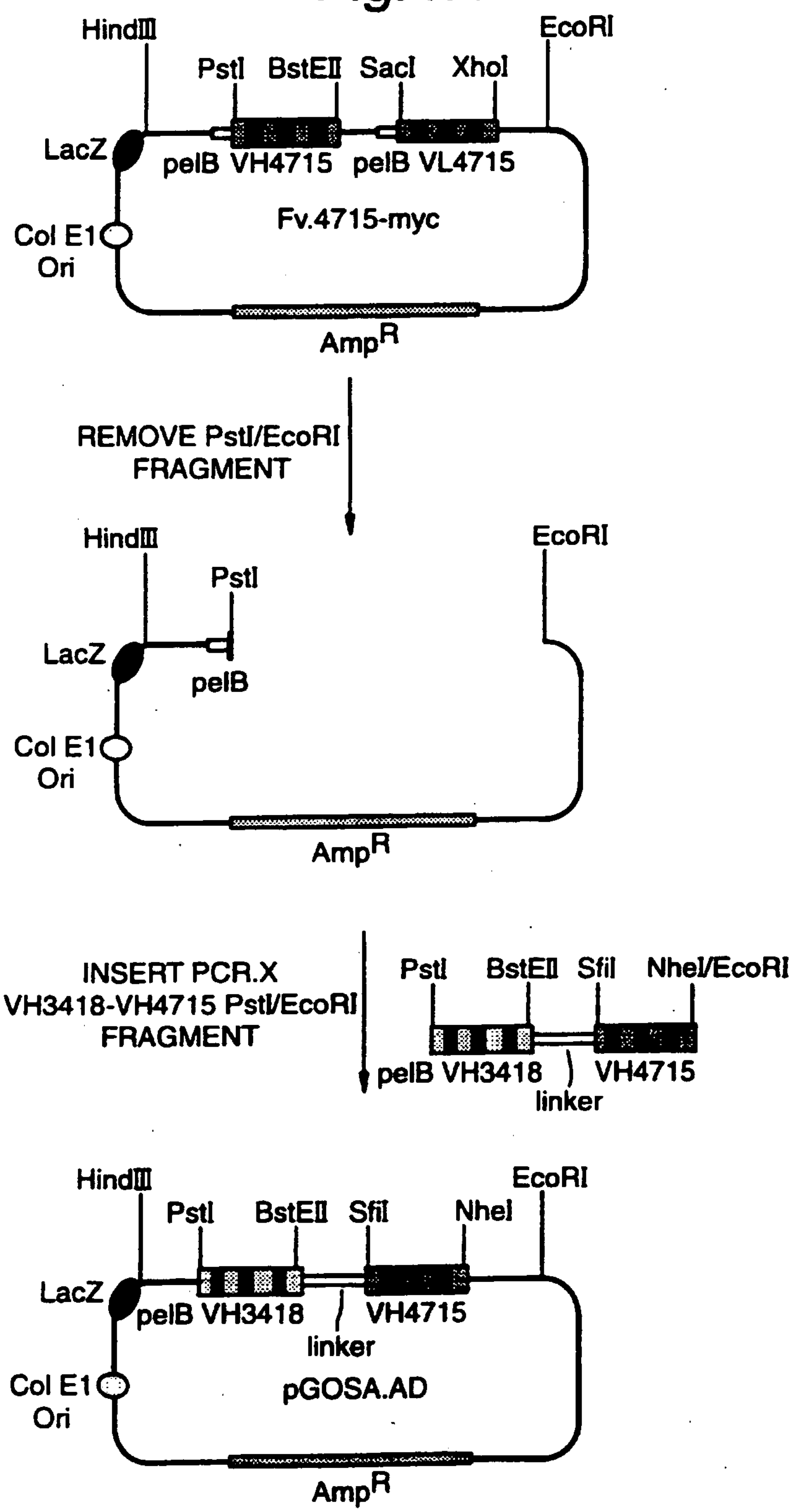

HindIII
PstI
BstEII
NheI
SacI
XhoI
SalI
NotI
EcoRI
LacZ
pelB VH3418 VH4715
pelB VL3418
VL4715
linker
Col E1
Ori
pGOSA.V
AmpR
1) REMOVE NheI/EcoRI VL3418-linker-VL4715
FRAGMENT FROM VECTOR
2) FILL IN ENDS WITH KLENOW DNA POLYMERASE
HindIII
PstI
BstEII
Blunt
Blunt
LacZ
pelB
VH3418 VH4715
Col E1
Ori
AmpR
CLOSE VECTOR
HindIII
PstI
BstEII
LacZ
pelB VH3418 VH4715
Col E1
Ori
pGOSA.AC
AmpR

Fig.47.


HindIII
EcoRI
PstI
BstEII
SacI
XhoI
LacZ
pelB VH4715
pelB VL4715
Fv.4715-myc
Col E1
Ori
AmpR
REMOVE PstI/EcoRI
FRAGMENT
HindIII
EcoRI
PstI
LacZ
pelB
Col E1
Ori
AmpR
INSERT PCR.X
VH3418-VH4715 PstI/EcoRI
FRAGMENT
PstI
BstEII
SfiI
NheI/EcoRI
pelB VH3418
VH4715
linker
HindIII
EcoRI
PstI
BstEII
SfiI
NheI
LacZ
pelB VH3418
VH4715
linker
Col E1
Ori
pGOSA.AD
AmpR

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 96/03605

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K16/46 C07K16/00 A61K39/395 G01N33/577

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,O	BIOCHEMICAL SOCIETY TRANSACTIONS, vol. 23, no. 4, 18 - 21 July 1995, LONDON, GB, pages 1067-1073, XP000565752 M. VERHOEYEN ET AL.: "Antibody fragments for controlled delivery of therapeutic agents." see the whole document ---	1-14

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 December 1996	07.01.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Nooij, F

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/EP 96/03605

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA, vol. 90, no. 14, 15 July 1993, WASHINGTON, DC, USA, pages 6444-6448, XP002021302 P. HOLLIGER ET AL.: ""Diabodies": Small bivalent and bispecific antibody fragments." cited in the application see the whole document ---	1-14
A	THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 269, no. 1, 7 January 1994, BALTIMORE, MD, USA, pages 199-206, XP002021303 W. MALLENDER ET AL.: "Construction, expression, and activity of a bivalent bispecific single-chain antibody." cited in the application see abstract see figures 1,2 ---	1-14
A	WO 94 13806 A (THE DOW CHEMICAL COMPANY) 23 June 1994 cited in the application see figure 1 ---	1-14
A	WO 94 13804 A (CAMBRIDGE ANTIBODY TECHNOLOGY LTD. ET AL.) 23 June 1994 cited in the application see page 31, line 10 - line 12 see figure 1 ---	1-14
A	WO 93 11161 A (ENZON, INC.) 10 June 1993 cited in the application see page 22, line 1 - line 10 see claims -----	1-14

1

INTERNATIONAL SEARCH REPORT

.iformation on patent family members

Inte- nal Application No
PCı/EP 96/03605

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9413806	23-06-94	AU-A-	5747794	04-07-94
		CA-A-	2117477	23-06-94
		EP-A-	0628078	14-12-94
		JP-T-	7503622	20-04-95
WO-A-9413804	23-06-94	AU-A-	5654894	04-07-94
		CA-A-	2150262	23-06-94
		EP-A-	0672142	20-09-95
		JP-T-	8504100	07-05-96
		AU-A-	7621494	10-04-95
		CA-A-	2169620	30-03-95
		EP-A-	0720624	10-07-96
		WO-A-	9508577	30-03-95
WO-A-9311161	10-06-93	AU-A-	3178993	28-06-93
		CA-A-	2122732	10-06-93
		EP-A-	0617706	05-10-94
		JP-T-	7501451	16-02-95